IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 5, 2019, accepted August 20, 2019, date of publication August 26, 2019, date of current version September 6, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2937326*

# Recognition Method of Green Pepper in Greenhouse Based on Least-Squares Support Vector Machine Optimized by the Improved Particle Swarm Optimization

**WEI JI[1], GUANGYU CHEN[1], BO XU[1], XIANGLI MENG[1], AND DEAN ZHAO[1,2]**

[1]School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China
[2]Key Laboratory of Facility Agriculture Measurement and Control Technology and Equipment of Machinery Industry, Jiangsu University, Zhenjiang 212013, China

Corresponding author: Wei Ji (jwhxb@163.com)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61703186 and Grant 61973141, and in part by the project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

**ABSTRACT** In the green pepper harvesting robot, the color of green pepper is similar to that of leaves, which makes it difficult to recognize the green pepper target. In order to solve this problem, a green pepper recognition method based on least-squares support vector machine optimized by the improved particle swarm optimization (IPSO-LSSVM) is proposed in this paper. Firstly, the green pepper images are segmented by K-Means method under the Lab color space, and the segmentation images of the target and background are obtained. The processed green pepper image was divided into training and testing samples. Then, the shape and texture features of green pepper targets are extracted separately from the training sample using the hu invariant moment and Tamura texture feature. Meanwhile, in order to reduce the complexity of data calculations and improve the efficiency, the extracted feature vectors are normalized. The feature vector is used as the input eigenvector of the least-squares support vector machine (LSSVM). The particle swarm optimization algorithm is used to obtain the optimal regularization parameter and the kernel function width. In order to maintain the particle activity, the mutation strategy is introduced to improve the particle swarm optimization algorithm. The experimental results show that the recognition rate of IPSO-LSSVM is higher than that of other methods, and the recognition accuracy is 89.04%. It could meet the requirements of green pepper identification.

**INDEX TERMS** K-means segmentation, feature extraction, PSO, harvesting robot, LSSVM.

## I. INTRODUCTION

In the greenhouse environment, the harvesting robot is the development trend of modern agriculture [1]–[3], and target recognition is an important step for successful picking. The speed and accuracy of the target recognition have a great influence on the working efficiency and time of the robot. A stable target recognition system allows the robot to work for a long time, compressing labor costs and increasing production efficiency. For green peppers, because the color of fruits is similar to leaves, it is very difficult to segment and recognize green pepper only according to its color feature compared with the large color difference between the target and background. Therefore, the development of a good visual recognition system picking robot has great application value and practical significance for the development of green pepper property.

Robotic crop harvesting and the methods for detecting crops have been explored by several researchers. Vitzrabin and Edan [4] proposed an adaptive threshold algorithm based on sensor fusion to identify sweet peppers. The RGB image is segmented into a number of rectangular sub-images. Three thresholds are adaptively calculated and applied for each sub-image, and the morphological operation is combined with the depth sensor to reduce the error. Thereby, the detection of the sweet pepper is realized. However, the algorithm does

The associate editor coordinating the review of this manuscript and approving it for publication was Jagdish Chand Bansal.

not take the noise problem of the image into account. Bac *et al*. [5] proposed a green pepper classifier that is robust to environmental changes. The green pepper plants were divided into two parts: soft (leaf and petiole) and hard (stem and fruit) obstacles. The plant characteristics were obtained by multi-spectral camera, and then the plants were classified by the classification regression tree (CART) classifier. Due to the change of illumination, the different shapes of plant parts and non-plant objects (support wires, drippers, etc.) ultimately lead to the low classification performance of the classifier. McCool *et al*. [6] proposed a scheme for classifying multispectral image data of green pepper targets using conditional random field (CRF). The program incorporates a range of texture features of green pepper, including local binary pattern (LBP), directional gradient histogram (HOG) and sparse autoencoder (SAE) features. These features are entered into the CRF for training, and the target fruit is finally detected. The method has a high recognition rate for target green pepper and the false recognition rate is relatively high, and a suitable method has not been found to reduce the false recognition rate. Kitamura and Oka [7] used LED lights to illuminate green peppers and converted the acquired RGB image into HSI color space. The S and I components are used to identify the reflective areas and to limit the area near the reflective region. And then carried out the two value of H, S, I component to complete the recognition. The green pepper has achieved good results, but it could not guarantee the recognition accuracy by lighting and color components. Zhang *et al.* [8] designed a hybrid classifier based on a support vector machine classifier of normalized *g* component, H and S component in HSV color space, and a threshold classifier characterized by super green operator. Although the recognition of the fruit is completed, it is not perfect in the image processing. It does not fully consider the problem of noise in the complex background, and the misclassification problem exists in the classification of the target. Sun *et al.* [9] used the fuzzy set theory and manifold ranking algorithm (FSMR) to identify green apples. By using fuzzy set theory in HSI color space, enhancing the original image makes the apple target more prominent. Select a specific query node to sort the pixels in the whole image, and finally identify the apple. However, this method requires further study due to limitations in identifying objects. He *et al.* [10] proposed a green litchi recognition method based on improved LDA classifier. The convolution kernel of the sample pixel is trained by extracting convolution characteristics. The pixels are classified by integrating multiple LDA classifiers with Ada Boost method, and litchi is identified by Hough transform circle detection method. However, the recognition rate of this method is low and the method needs to be improved.

In this paper, the green pepper in the greenhouse is taken as the research object. Firstly, the green pepper images are segmented by K-Means method under the Lab color space, and the segmentation images of the target and background are obtained. The processed green pepper image was divided into training and testing samples. Then, the extracted shape and texture features are used as input vectors of the LSSVM. The particle swarm optimization algorithm is used to obtain the optimal regularization parameter and the kernel function width. In order to maintain the particle activity, the mutation strategy is introduced to improve the particle swarm optimization algorithm. The least-squares support vector machine is used for recognition. Finally, the effectiveness of the algorithm is verified.

## II. RELATED WORKS

### A. VISION SYSTEM SETUP AND IMAGE ACQUISITION

The green pepper image is collected with the UniFly M088 camera on the harvesting robot, and 100 images are randomly collected. The image are 640 pixels by 480 pixels. The computer used for image processing platform is: Intel (R) Core (TM) i5-3230MCPU@2.60GHz, and 4G memory. The software platform designed by opencv3.0 is combined in Visual Studio 2013 to perform simulation verification of the experiment.

### B. IMAGE PROCESSING

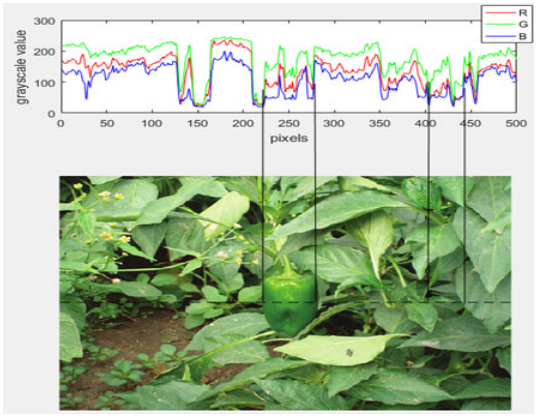
The green pepper background is complex and diverse, including the background elements (such as branches, fruit stalks, green leaves and so on) similar to fruit color. The growth status of green peppers included separation, adjacency, and the partial shade of the fruit.

Fig. 1 shows the R, G, and B components of the collected green pepper images. The three curves represent the R, G, and B components of the pixels at the horizontal black line in the image [11]. The green component of pepper is very close to the leaves, which makes it difficult to distinguish in the process of recognition [12].

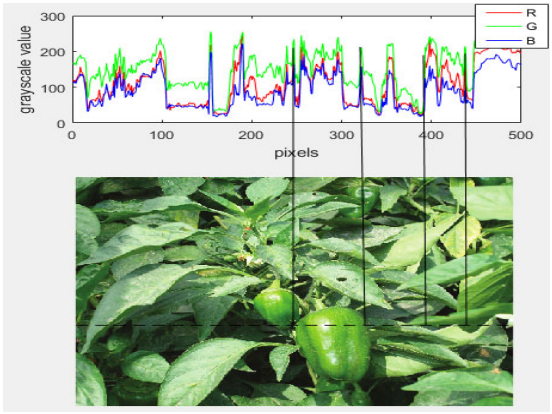
The choice of color space often determines the quality of image segmentation. The common color spaces are RGB, XYZ, Lab and so on. The RGB color space can represent most of the colors, but the correlation among each component is too strong, so it is not suitable for image segmentation directly [13]. However, the main disadvantage of the XYZ color space is non-uniformity. In different areas of chromaticity and brightness, the resolution of human is inconsistent. So the XYZ color space should not be used directly for image segmentation. The Lab color space has the advantage of separating color information and making human perception of brightness linear. Therefore, the RGB image is transformed into Lab color space for segmentation in this paper [14].

### C. IMAGE SEGMENTATION

The green pepper samples in the collected images are not uniform in color due to the interference of uneven illumination and noise. The target often presents discontinuities and irregularities, and ordinary methods are difficult to meet the requirements [15], [16]. The K-Means algorithm is a clustering algorithm based on distance measure. Firstly, the initial clustering centers are determined according to specific rules,




(a)




(b)

**FIGURE 1.** R, G and B component polygraphs.

but the initial classification is not necessarily reasonable. Then, the clustering centers are adjusted according to the minimum distance or the least square principle in the group until the clustering centers are no longer changed. At this time, the relatively reasonable classification results will be output. K-Means maintains scalability and efficiency in dealing with large data. It is more efficient than other algorithms. The K-Means clustering algorithm takes K as a parameter, which is to divide the $m$ samples into $k$ different classes, so that the samples in the class have stronger similarity. The average difference between clusters is calculated according to the $u_k$.

$$E_k = \sum_{i}^{m} (x_{ik} - u_k)^2 \tag{1}$$

where $x_{ik}$ is the *ith* sample in the *Kth* cluster. By iterating, the sum of squared errors of the aggregated K classes is minimized. That is, the distance within each class is as small as possible.

The steps of image segmentation of green pepper based on K-Means clustering algorithm are as follows in Fig. 2.

The collected green pepper images are converted to Lab color space, and K-Means clustering is performed by calculating the L, a, and b values of each pixel. Because of the texture and shape difference of peppers and leaves,

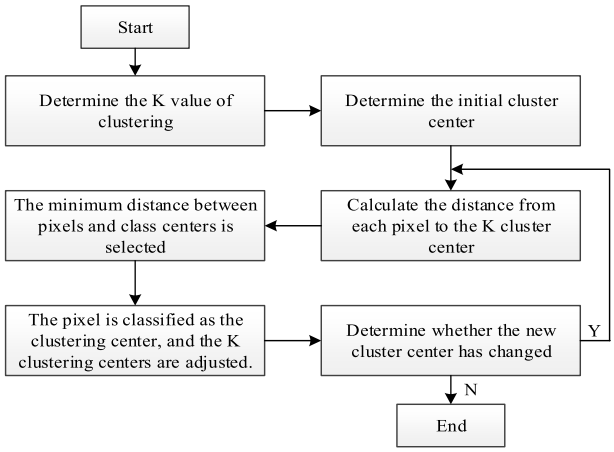



**FIGURE 2.** Basic steps of K-Means segmentation.

the value of K is 2, which represent peppers and leaves. In this paper, the Euclidean distance is adopted in the sample distance, and the minimum distance principle is adopted in the clustering criterion. The segmentation diagram is shown in Fig. 3.

Based on the K-Means method, the image is divided into two parts of green pepper and leaves. However, during the segmentation process, there will be noise in the image. According to the basic morphological characteristics of green pepper, the segmented image is opened and closed by using a $5\times5$ element matrix to remove part of the noise [17]. It can be seen from the figure that the segmentation result is well, reaching the expectation, which is convenient for the next step of recognition processing.

### D. IMAGE FEATURE EXTRACTION

The choice of features is crucial for the extraction of targets [18]. The typical characteristics can effectively represent the target. The quality of the extracted features is directly related to the accuracy of the recognition. In this paper, the shape features can intuitively distinguish different targets. The Hu invariant moments have translation, rotation and scaling invariant, which can better express the characteristics of the target. Texture features describe the surface properties of the target corresponding to the image area. Texture feature is different from color feature, it is not based on pixel feature. Texture features require statistical calculations in regions that contain multiple pixels. It will contain more target details after combining shape features with texture features. So that the image can be recognized accurately [19].

#### 1) SHAPE FEATURE EXTRACTION

Green pepper and leaves are similar in color, so the recognition by color characteristics cannot ensure the accuracy [20]. In this section, the shape feature is extracted. According to the characteristics of the segmentation image, the shape feature of the image is extracted by the Hu-invariant moment method, which is translation, rotation and scaling invariant [21].

(a) The original image (b) The segmented image (c) The labeled connected regions image (d) The segmented leaf image

**FIGURE 3.** K-Means segmentation results of green pepper images.

Hu-invariant moment is used to extract shape features from target connected regions. A pixel on the connected region is $(x, y)$, and the corresponding gray value is $f(x, y)$. Then the (p + q) order moment of the whole connected region is:

$$m_{pq} = \sum_{x} \sum_{y} f(x, y) x^p y^q \tag{2}$$

The image of centroid coordinates are:

$$x_0 = {}^{m_{10}}/_{m_{oo}}, \quad y_0 = {}^{m_{ol}/m_{oo}} \tag{3}$$

The center distance is calculated as follows:

$$\mu_{pq} = \sum_{x} \sum_{y} (x - x_0)^p (y - y_0)^q f(x, y) \tag{4}$$

From the above equation, the center moment is position independent and translation invariant. In order to make the center moment also have scale and rotation invariant, we need to normalize the Hu-invariant moment, denoted as $\eta_{pq}$.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{1+(p+q)/2}} (p + q = 2, 3, 4 \ldots) \tag{5}$$

Therefore, Hu-invariant moments can be obtained as follows:

$$\begin{aligned}
a_1 &= \eta_{02} + \eta_{20} \\
a_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
a_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{13})^2 \\
a_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
a_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{21})^2 - (\eta_{21} + \eta_{03})^2\right] \\
a_6 &= (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
a_7 &= (3\eta_{21} - \eta_{03})(\eta_{10} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{133})^2\right]
\end{aligned} \tag{6}$$

The Hu-invariant moment eigenvalues of five target regions and leaves are selected arbitrarily. As can be seen from Fig. 4, the whole moments of invariant moments show a decreasing trend. At the same time, there is a significant difference in the moment invariants between green peppers and leaves. Therefore, it is feasible to obtain the shape features of the target by using invariant moments.

### 2) TEXTURE FEATURE EXTRACTION

The texture feature is a visual feature that reflects the homogeneity of the image without relying on color and brightness [22]. It can express the local pattern and intrinsic characteristics of the image. Including the arrangement of the surface structure of the object and important information related to the surrounding environment. The Tamura texture feature is a well-proven algorithm for texture metrics. It is more intuitive than the grayscale co-occurrence matrix, and it has more visual effects. In general, the three components of coarseness, contrast, and directionality have important applications in image recognition [23], [24].

Coarseness reflects a quantity of grain in the texture. Firstly, we need to calculate the average gray value of the pixels in the active area of the $2^k \times 2^k$ pixel size in the image.

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{2^{k-1}-1} g(i, j)/2^{2k} \tag{7}$$

where $k = 0, 1, 2 \ldots 5$, $g(i, j)$ is the pixel gray value at $(i, j)$.

The average gray values of the $2^k \times 2^k$ pixels in the horizontal and vertical directions are calculated respectively.

$$\begin{aligned}
E_{k,h}(x, y) &= \left|A_k\left(x + 2^{k-1}, y\right) - A_k\left(x - 2^{k-1}, y\right)\right| \\
E_{k,y}(x, y) &= \left|A_k\left(x, y + 2^{k-1}\right) - A_k\left(x, y - 2^{k-1}\right)\right|
\end{aligned} \tag{8}$$

From the above equation, each pixel will get a group E values, and the best size of each pixel will be $S_{best}(x, y) = 2k$.The coarseness of this image is obtained by averaging the $S_{best}$ of the entire image pixel.

$$F_{crs} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} S_{best}(i, j) \tag{9}$$

Contrast is the level description of the luminance between the brightest and darkest regions in the image. The greater the difference between objects, the greater the contrast. The contrast is defined by $\alpha_4 = \mu_4/\sigma^4$, where $\mu_4$ is the fourth

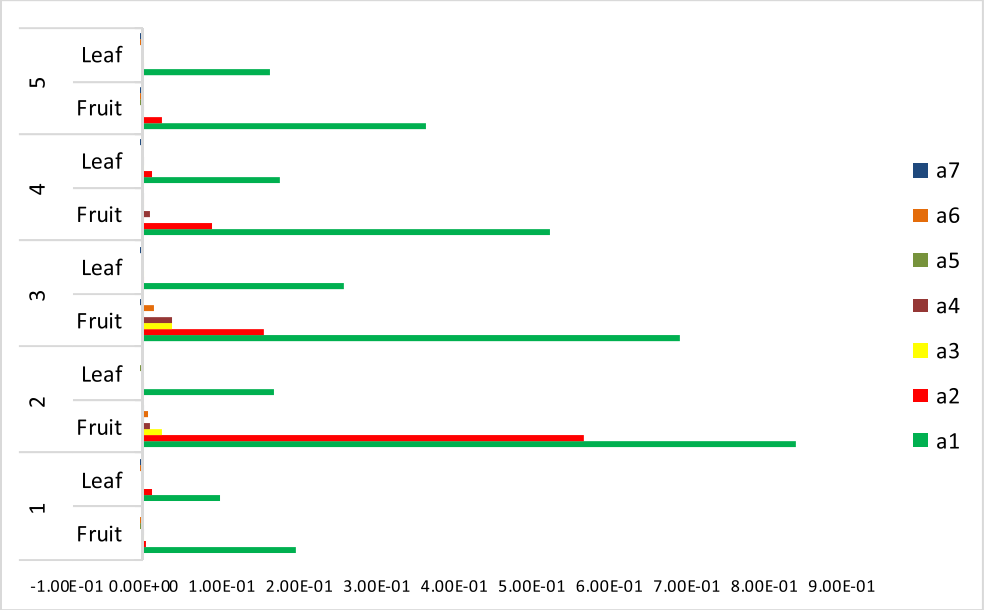



**FIGURE 4.** Hu-invariant moment image.

moment and $\sigma^2$ is the variance. The formula for measuring contrast is as follows:

$$F_{con} = \frac{\sigma}{\alpha_4^{1/4}} \tag{10}$$

The value can express the global variable of the contrast in the entire image.

Directionality describes the distribution or concentration in certain directions through the global characteristics of texture regions. The degree of direction is calculate the gradient vector at each pixel [25].

Divide the angle of $0 \sim \pi$ into 16 parts. The maximum value of each interval is identified as $\varphi$. A function $n = HD(\varphi)$ describing the distribution of $\theta$ angles in an image is created. The dependent variable is the number of gradient vectors in every bisection interval. The independent variable $\varphi_p$ is the amount that gives the maximum value to the function $n$. The directivity of the image is obtained by calculating the sharpness of the peak value in the histogram of $\theta$.

$$F_{dir} = \sum_{p}^{n_p} \sum_{\varphi \in w_p} (\varphi - \varphi_p)^2 H_D(\varphi) \tag{11}$$

where $p$ is the peak value and $n_p$ is all the peaks in the histogram. $w_p$ is all the discrete regions included in the peak $p$, and $\varphi_p$ is the center of the peak.

For randomly selecting five green pepper images, the coarseness, contrast and directionality of green pepper and leaves are simulated. The results are shown in Fig. 5.

It can be seen from Fig. 5 that green peppers and leaves have significant differences in coarseness, contrast and directionality. Therefore, it is effective for the texture features of the classifier to select these three feature vectors.

### E. DATA NORMALIZATION

The acquired shape and texture features data have different dimensions and magnitudes. Therefore, it is necessary to convert data to ensure the effective use of data to achieve accurate identification. The original data matrix is converted to a new row or column of elements that do not rely on the dimension and magnitude in accordance with some rules of operation [26].

The initial sample size of the data is $n$ and the feature vector is $m$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,m} \\ a_{21} & a_{22} & \cdots & a_{2\,m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}_{n \times m} \tag{12}$$

There is a variety of data transformation methods for the unified dimension and magnitude of the feature of each column. This paper adopts the data normalization process:

$$x_{ij} = \frac{x_{ij} - \overline{x}_j}{\max\left(x_j\right) - \min\left(x_j\right)} \tag{13}$$

where $x_j = \sum_{i=1}^{n} x_{ij}$, $i = 1, 2 \ldots n, j = 1, 2 \ldots m$.

The average value of each element in a column is calculated, and the numerator is the original element subtracts the average value. At the same time, the denominator is the maximum value in the column subtracts the minimum value, and the normalized data range is $[-1, 1]$.

## III. RECOGNITION METHOD

The seven shapes and three texture eigenvalues were obtained according to the shape and texture features of the target connection region, and the data was normalized. The segmented

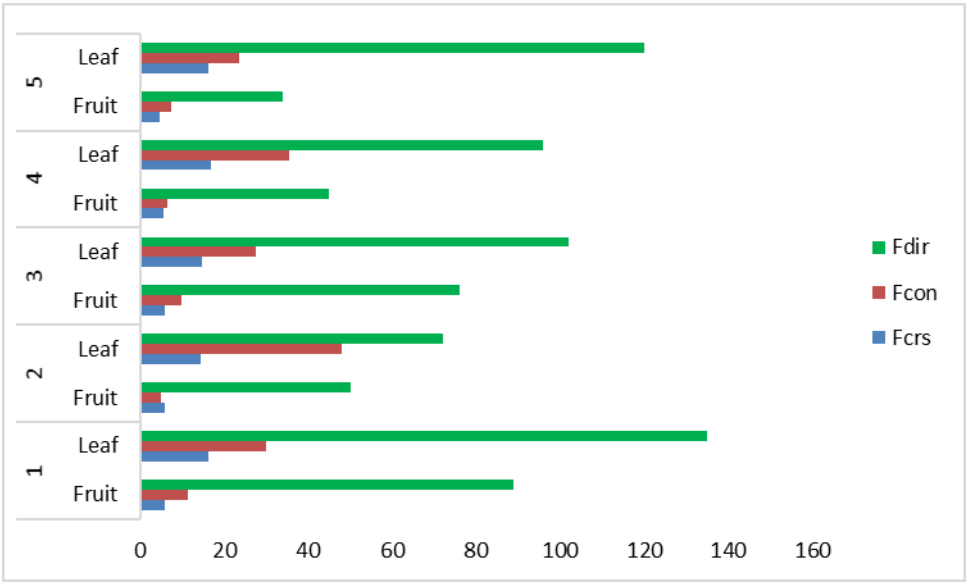



**FIGURE 5.** Texture feature results.

image was divided into training and testing samples in order to accurately identify the green pepper in the image. Then the least-squares support vector machine based on improved particle swarm optimization was used to train the features of training samples, and the recognition algorithm was obtained. Finally, the testing sample were identified and verified, and the results were analyzed.

### A. THE LEAST-SQUARES SUPPORT VECTOR MACHINE

In this paper, the texture and shape vectors are extracted for training the SVM classifier. The SVM maps input samples from low-dimensional to high-dimensional feature space by using the non-linear mapping function [27], [28]. The SVM constructs the optimal classification plane to solve the linear inseparability problem of the original space in the high dimensional feature space [29]. However, the data cannot be separated by a hyperplane because texture vector data is linear non-separable data [30]. Therefore, the problem is solved by following the principle of minimizing structural risks. The regularization parameter ($\gamma$) describes the generalization ability of the function. At the same time, the kernel function is adopted to transform the data nonlinearly to higher dimension in which the hyperplane is found to separate the data. In this experiment, the Gaussian kernel is adopted.

$$k(x, x_i) = \exp[-\frac{\|x - x_i\|^2}{2\sigma^2}] \tag{14}$$

where $k$ is the kernel function, $x$ and $x_i$ are vectors of training data, $\sigma$ is the width of the kernel function [31].

According to the working principle of LSSVM, the kernel function width ($\sigma$) and the regularization parameter ($\gamma$) determine the learning and generalization capabilities. The value of the kernel function width has an effect on the distribution complexity of the sample in the feature space. When the value of the kernel function width increases, the samples are correctly classified. But the value of the kernel function width is too large, it will cause over fitting and weaken the generalization ability. If the regularization parameter is too small, it will be increased the fitting error. Therefore, the value of the kernel function width and the regularization parameter will have a direct impact on the performance of the LSSVM classification algorithm. At the same time, two parameters are selected to get the best combination of parameters.

### B. IMPROVED LSSVM PARAMETERS OPTIMIZED BY PARTICLE SWARM OPTIMIZATION

In support vector machine, the main factors that affect the result are regularization parameter and kernel function width [32]. In this section, PSO consists of a swarm of particles that search for the best position with respect to the corresponding best solution for an optimization problem in the virtual search space [33]. The optimal solution of two parameters is found by improving the ability of global optimization of particle swarm optimization, and the best classification results are achieved by solving the best hyper plane [34]. The optimized method does not improve the quality of the solution after repeated iterations of particles. The current position of the particles has changed, so it does not need to fall into the local optimal solution [35], [36]. The flow chart of the IPSO-LSSVM is shown in Fig.6. And the detailed description is as follow:

1. The continuous values of regularization parameters and kernel function widths are represented by two-dimensional particle coded individuals.

2. Initialize the population randomly and specify parameters. Such as the ceiling, minimum, initial velocity, population size and number of iterations.

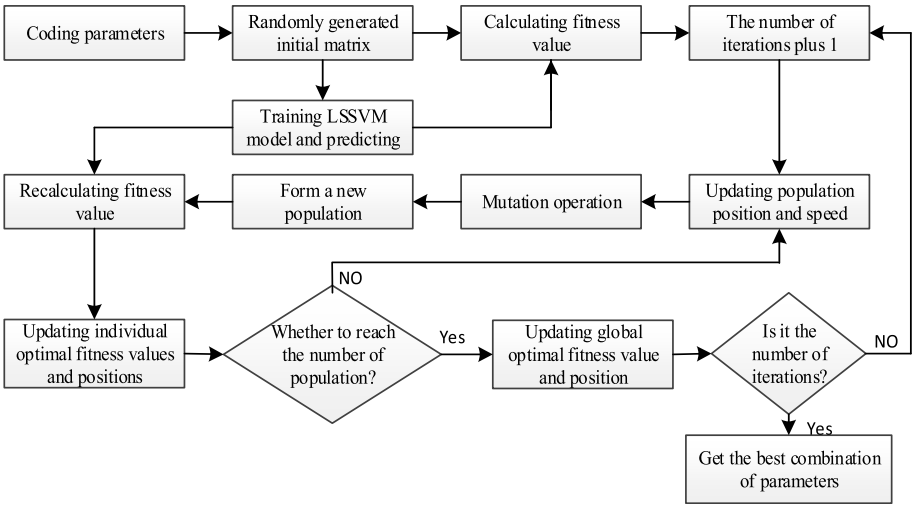



**FIGURE 6.** The flow chart of the IPSO-LSSVM algorithm.

3. The LS-SVM model is trained by the feature set of step 2.

4. In general, when the ACC value is large, the performance of the classifier is better. When the number of support vectors is small, the generalization ability of the SVM is strong. Considering the design of their objective function, the fitness value is calculated according to the objective function.

$$\begin{cases} f_1 = \dfrac{\sum_{i=1}^{K} Test\_Accuracy_i}{K} \\ f_2 = (1 - \dfrac{nsv}{m}) \\ f = \alpha \times f_1 + \beta \times f_2 \end{cases} \tag{15}$$

where $f_1$ denotes the classification accuracy value of the LS-SVM classification model. In $f_2$, $nsv$ and $m$ denote the number of support vectors and the total number of samples in the training sample. The objective function is to weight the sum of two sub-goals. For the function $f$, $\alpha$ denotes the weight of the LS-SVM model accuracy and $\beta$ is the weight of the LS-SVM number. Represent $\alpha$ and $\beta$ as follows.

$$\begin{aligned} \alpha &= (\alpha_1 - \alpha_2)\frac{t}{t_{\max}} + \alpha_2 \\ \beta &= (\beta_1 - \beta_2)\frac{t}{t_{\max}} + \beta_2 \end{aligned} \tag{16}$$

where $\alpha$ and $\beta$ satisfy $\alpha_1 + \beta_1 = 1$ and $\alpha_2 + \beta_2 = 1$. After obtaining the fitness value, the global optimal fitness value is set to *gfit* and the individual optimal fitness value is set to *pfit*. The position of the global optimal particle is set to *gbest* and the position of the individual's history is set to *pbest*.

5. The number of iterations was increased.

6. The number of populations was increased. The position and velocity of $\gamma$ and $\sigma$ for each particle will be updated simultaneously.

7. In order to improve the performance of particle swarm algorithm, a mutation strategy was introduced to improve it.

$$\begin{aligned} &\text{If}(count(gbest) \geq 5) \\ &\text{then } gbest = 0; \quad \sum_{i=1}^{n} ft_i = 0. \end{aligned} \tag{17}$$

where the function of *count* indicates that the *gbest* value remains unchanged for several times, and $ft_i$ represents the selected feature number. The entire strategy represents that if the *gbest* value remains unchanged for five consecutive times, it is considered to be trapped in a local extremum.

The LS-SVM model is trained by the feature of step 6 and then calculate the fitness value of each particle.

8. The current fitness value with *pfit* in memory were compared to update the individual optimal fitness value and position. If the fitness value is smaller than that in memory, the *pfit* and the *pbest* in the record are kept. Otherwise, the two are replaced by the fitness values and the position of the particles at this time.

9. If the number of particles is equal to the maximum size, the steps will continue. Otherwise, the step will go to step 6.

10. If the best *pfit* is smaller than the *gfit* in the history record store, it will keep the global optimal fitness value *gfit* and the global optimal fitness value *gbest*. Otherwise, it will be replaced with historical record values.

11. If the number of iterations reaches the limit, it will continue to execute. Otherwise, the step will jump to step 5 to continue the iteration.

12. The optimal parameter regularization parameters and the kernel function width were obtained from the global optimal position.

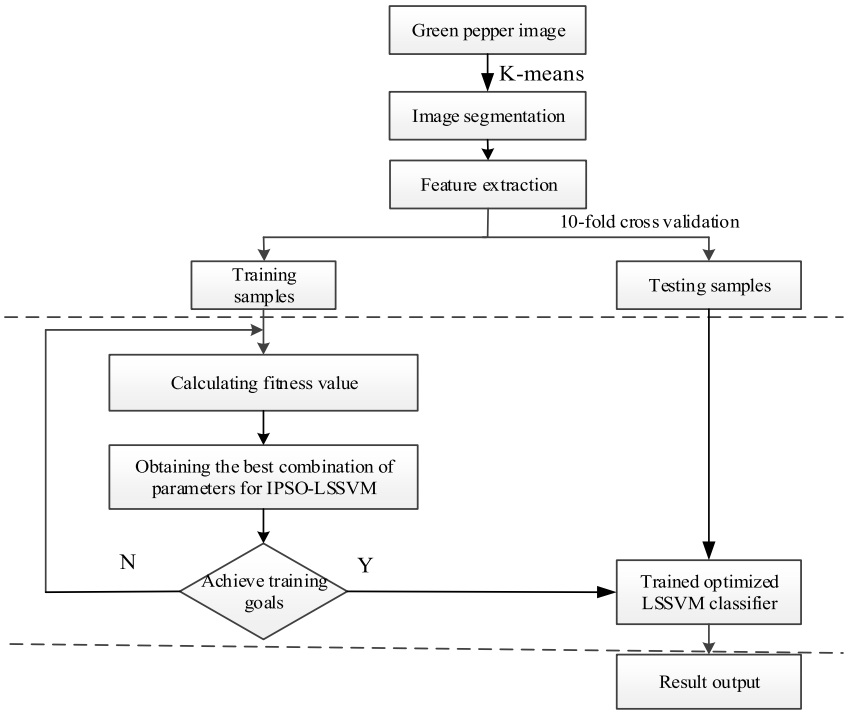



**FIGURE 7.** The flow chart of green pepper recognition algorithm.

### C. GREEN PEPPER RECOGNITION PROCESS BASED ON IPSO-LSSVM

The algorithm of green pepper recognition is divided into the following steps. Firstly, we need to divide the image by K-Means method, and extract the shape and texture feature after the denoising [37]. Then we construct the training sample and using the training data to select the least squares support vector machine parameters. In the end, the LSSVM with superior performance is obtained by analyzing the operation parameter selection method of LSSVM. The flow chart of green pepper recognition algorithm is shown in Fig. 7.

## IV. RESULTS AND ANALYSIS

### A. SEGMENTATION RESULTS

The K-means image segmentation was perfomed on 100 green pepper images, and the effective connected regions were preserved after filtering and denoising. Fig. 8 enumerates the connected region graph after image processing.

The processed of green pepper graph in Fig. 8 was colored, which largely preserved the detail information. Because of studying the recognition of the near-color targets, the K-Means method was used to segment the target and background, which effectively described the characteristics of green pepper samples.

### B. FEATURE EXTRACTION RESULTS

The shape and texture features are extracted respectively from the connected regions of the 100 processed images. Each set of data contains seven shape and three texture eigenvalues.

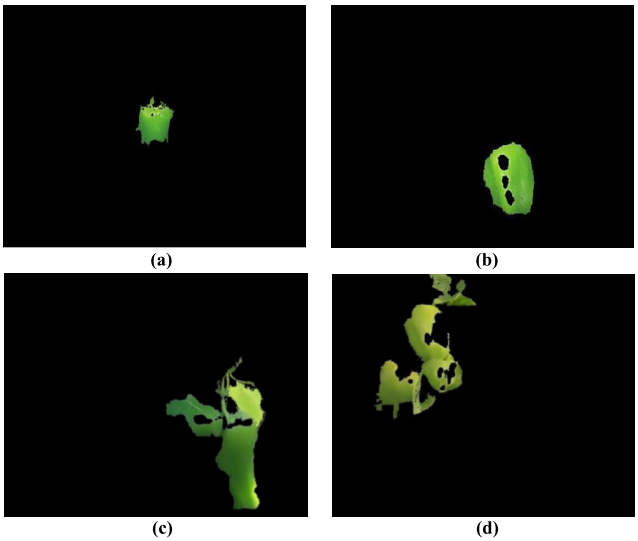



**FIGURE 8.** The segmentation image of green pepper.

We list 10 sets of eigenvalues of the corresponding samples, as shown in Table 1.

### C. ALGORITHM TRAINING RESULTS

The extracted shape and texture feature vectors are used as the feature vectors to be input. The PSO-LSSVM and IPSO-LSSVM are compared to verify the effectiveness of the improved particle swarm optimization based support vector machine model. In the experiment, the parameters $C_1$ and $C_2$ of PSO algorithm are set to 1.6 and 1.8 respectively.

**TABLE 1.** The eigenvalues of 10 groups of green pepper images.

| Values / Groups | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ | $F_{crs}$ | $F_{con}$ | $F_{dir}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.9850e-001 | 2.8939e-003 | 2.4735e-004 | 2.4723e-006 | -4.6769e-011 | -6.7696e-008 | 3.9376e-011 | 5.9009 | 11.2675 | 89 |
| 2 | 2.9112e-001 | 4.9281e-003 | 2.1950e-003 | 5.0259e-003 | 1.3582e-005 | 3.0002e-004 | 9.7055e-006 | 5.8841 | 5.3148 | 51 |
| 3 | 6.9517e-001 | 1.5685e-001 | 3.8399e-002 | 3.6885e-002 | 1.3773e-003 | 1.4066e-002 | -1.7349e-004 | 5.7704 | 5.0353 | 50 |
| 4 | 5.2678e-001 | 9.0291e-002 | 9.4638e-004 | 8.5183e-003 | 2.3494e-005 | 2.5373e-003 | 5.7434e-006 | 5.2995 | 5.2232 | 83 |
| 5 | 3.6577e-001 | 2.3523e-002 | 4.5493e-004 | 3.0687e-004 | -3.1640e-008 | -1.4391e-005 | -1.1021e-007 | 5.2771 | 5.4490 | 50 |
| 6 | 1.7665e-001 | 5.1762e-004 | 2.8747e-006 | 9.9341e-006 | -5.2279e-011 | 1.0865e-007 | 9.2278e-012 | 6.3401 | 6.2968 | 65 |
| 7 | 3.4895e-001 | 7.3522e-003 | 2.0842e-003 | 2.6273e-003 | 4.3168e-006 | 2.2492e-004 | -4.3774e-006 | 5.7869 | 9.6700 | 75 |
| 8 | 2.7624e-001 | 3.5215e-003 | 1.0746e-003 | 3.7919e-003 | 5.0556e-006 | 1.3375e-004 | -5.7473e-006 | 7.8186 | 6.8814 | 60 |
| 9 | 3.9340e-001 | 1.6230e-002 | 4.6329e-003 | 4.3045e-003 | -3.5532e-006 | -6.3998e-005 | -1.8891e-005 | 6.7974 | 8.6930 | 100 |
| 10 | 2.0092e-001 | 7.2395e-004 | 5.2574e-005 | 8.5188e-005 | -4.9027e-009 | -2.2464e-006 | 2.9095e-009 | 8.2821 | 10.5681 | 116 |

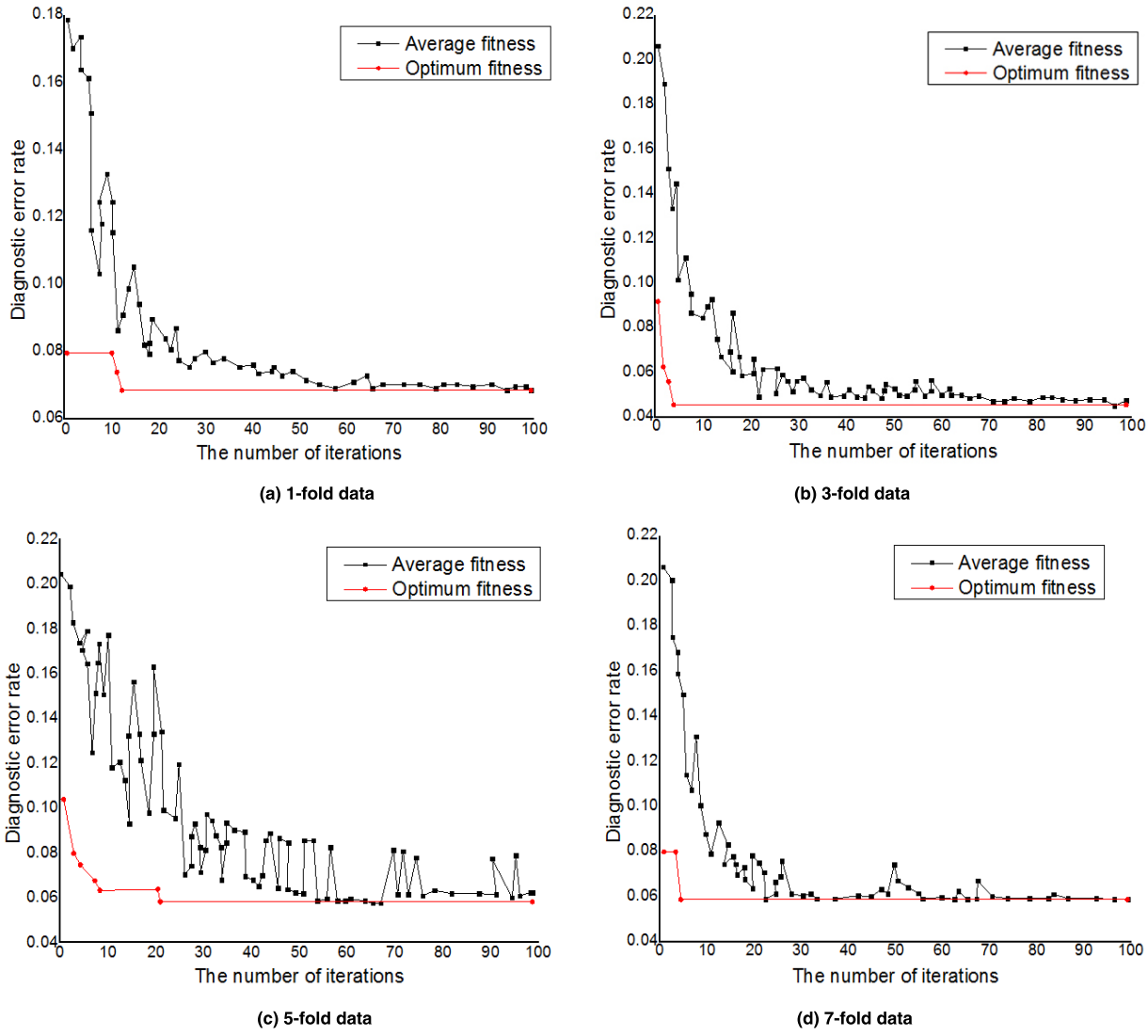



**FIGURE 9.** The recognition accuracy of PSO-LSSVM varies with the number of iterations.

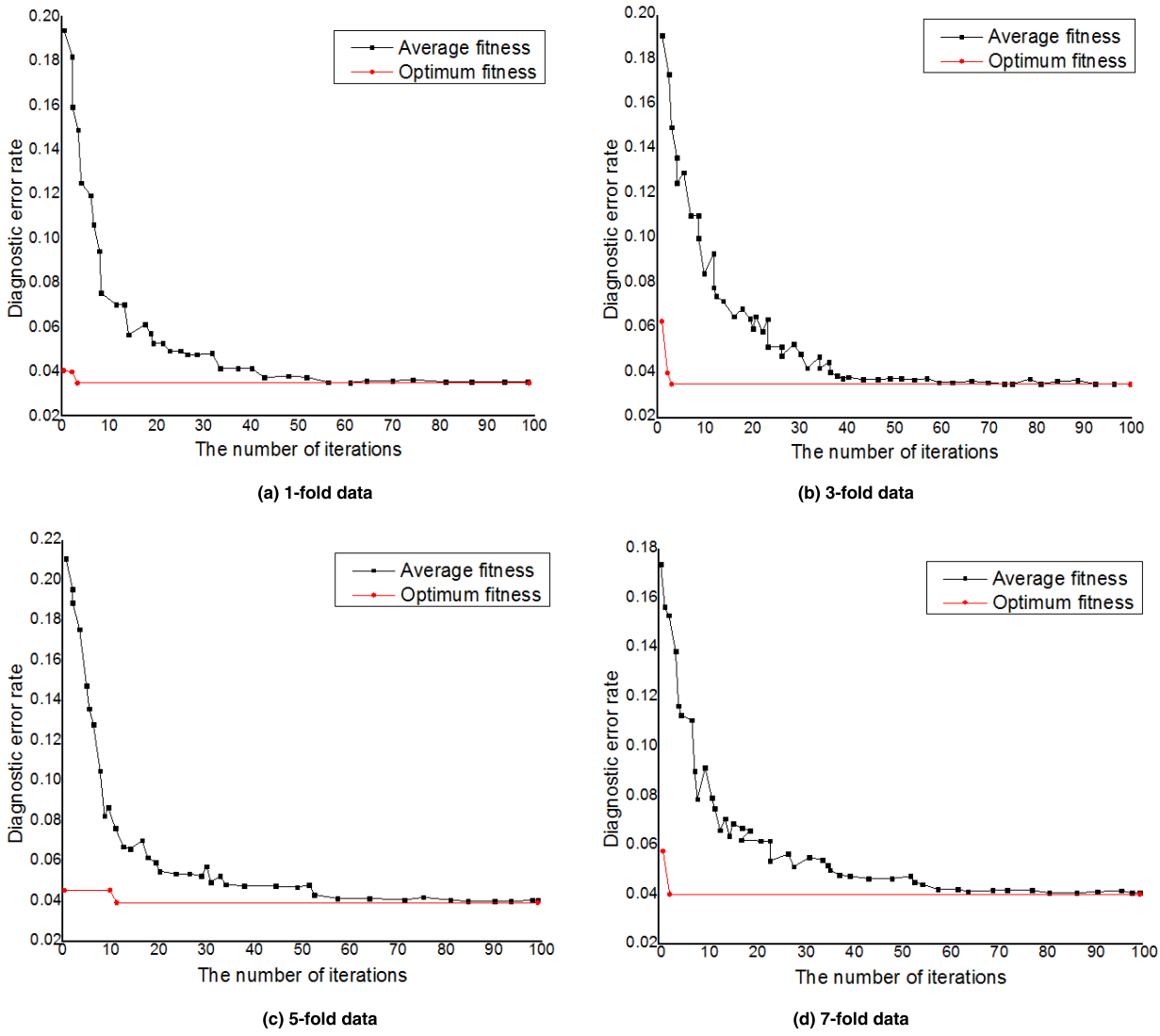



**FIGURE 10.** The recognition accuracy of IPSO-LSSVM varies with the number of iterations.

In the experiment, the performance of PSO-LSSVM is compared with IPSO-LSSVM. Fig. 9 is the process of the PSO-LSSVM recognition accuracy in the quadruple folding data in the 10-fold cross-validation (an algorithm for testing algorithm accuracy). From the figure, it can be observed that the number of iterations need to obtain the best fitness value on the data PSO-LSSVM algorithm is generally within 20 times, after which the accuracy no longer changes until the end of the iteration. This shows that the PSO can quickly converge to the optimal solution, so that the LSSVM obtains the optimal parameters so as to obtain the best recognition accuracy.

Fig. 10 is the process of the IPSO-LSSVM recognition accuracy in the quadruple folding data in the 10-fold cross-validation. It can be seen from the figure that the number of iterations required to obtain the best fitness value on the IPSO-LSSVM algorithm is generally around 10 times, and then the accuracy no longer changes until the end of the iteration.

Compared with Fig. 9, it can be seen in Fig. 10 that the IPSO can get higher solution and converges faster than the PSO algorithm. This shows that in order to maintain the particle activity, the method of mutation operation to the particle can prevent the particle from falling into the local optimal solution. The least squares support vector machine can obtain the best parameter, thus the higher recognition accuracy can be acquired.

## V. RECOGNITION RESULTS

Four algorithms are tested using 40 of the 100 green pepper pictures according to the correct detection rate, error recognition rate, leakage recognition rate and the average recognition time. Table 2 shows that the recognition results between four different algorithms. The experimental results show that the accuracy of the PSO-LSSVM can reach 82.2%. The accuracy of the IPSO-LSSVM can reach 89.04%, which is higher than 80.8% of CRF and 58.9% of CART respectively. It shows that this method is the most accurate in green pepper

**TABLE 2.** The recognition results of green pepper by different algorithms.

| Methods | The number of green peppers | Correct recognition | | Error recognition | | Leakage recognition | | Average recognition time (ms) |
|---|---|---|---|---|---|---|---|---|
| | | number | Recognition rate | number | Error Recognition rate | number | Leakage Recognition rate | |
| CART | 73 | 43 | 58.9% | 6 | 8.2% | 30 | 41.1% | 685 |
| CRF | 73 | 59 | 80.8% | 12 | 16.4% | 14 | 19.2% | 260 |
| PSO-LSSVM | 73 | 60 | 82.2% | 5 | 6.8% | 13 | 17.8% | 385 |
| IPSO-LSSVM | 73 | 65 | 89.04% | 3 | 4.1% | 8 | 10.9% | 320 |

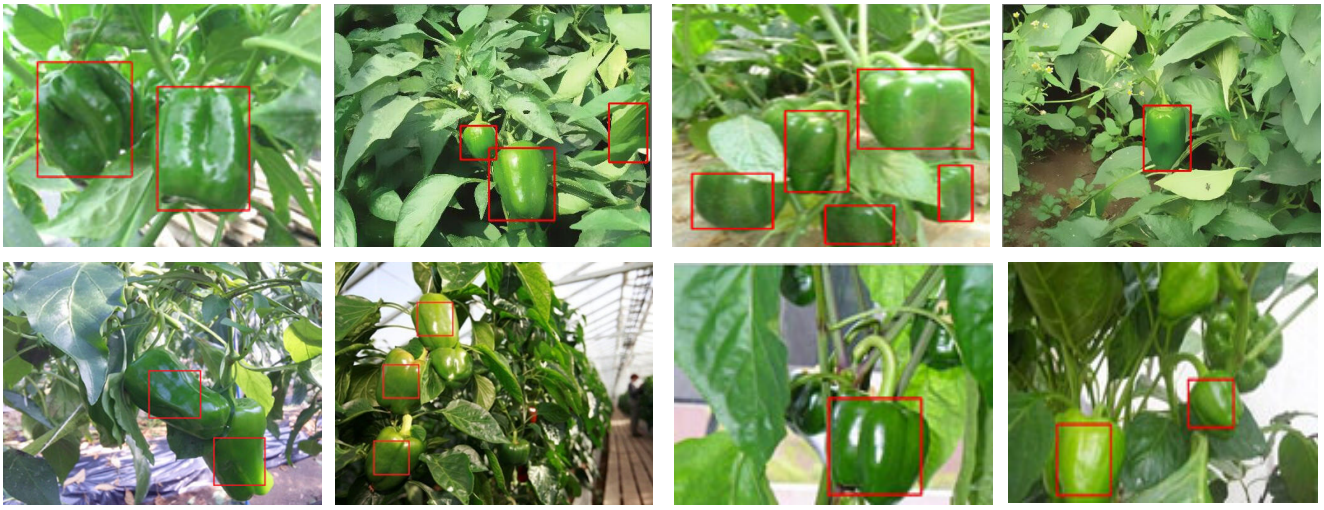


**FIGURE 11.** The recognition image of green pepper.

recognition and can obtain higher precision. For the average recognition time, the PSO-LSSVM algorithm is 385ms. The average recognition time of the IPSO-LSSVM algorithm is 320ms, which is lower than 260ms of CRF algorithm, but higher than 685ms of CART. However, the image acquisition does not need to be refreshed frequently in the process of harvesting, so the algorithm can basically meet the actual requirements.

It is obvious that the success rate of IPSO-LSSVM recognition is higher than that of PSO-LSSVM, CART and CRF. The error recognition rate and the leakage recognition rate are also lower than other three methods, which fully proves that the IPSO-LSSVM has better superiority. It can meet the requirements of accurate and efficient by the green pepper harvesting robot.

Fig. 11 shows that the recognition results of green pepper images using proposed method. It can be seen that green pepper is well detected. However, some similar leaves and obscured green peppers to be mistaken and missed, then further research and improvement of this algorithm are needed.

## VI. CONCLUSION

In this paper, according to the difficulty of green pepper recognition in near-color background, we propose an image recognition method based on improved particle swarm optimization (IPSO) least squares support vector machine (LSSVM). Based on the LSSVM for image recognition of green pepper. Firstly, the image of green pepper was acquired by UniFly M088 camera and image segmentation was performed by K-means algorithm, then the segmentation image was filtered. The processed green pepper image was divided into training and testing samples. Then, the extracted shape and texture features are used as input vectors of the LSSVM. The particle swarm optimization algorithm is used to obtain the optimal regularization parameter and the kernel function width. In order to maintain the particle activity, the mutation strategy is introduced to improve the particle swarm optimization algorithm. The experimental results show that the green pepper recognition model is obtained by the least squares support vector machine by improved particle swarm optimization algorithm. The training model is used to test the testing samples. The recognition rate of green pepper is 89.04%, and the average recognition time is 320ms. It meets the requirements of accuracy and time efficiency of harvesting robot for greenhouse green pepper recognizing. However, due to the high rate of leak recognition, the correct recognition rate of green pepper needs to be improved.

## REFERENCES


[1] T. Rath and M. Kawollek, "Robotic harvesting of Gerbera Jamesonii based on detection and three-dimensional modeling of cut flower pedicels," *Comput. Electron. Agricult.*, vol. 66, no. 1, pp. 85–92, 2009.

[2] X. Liu, W. Jia, C. Ruan, D. Zhao, Y. Gu, and W. Chen, "The recognition of apple fruits in plastic bags based on block classification," *Precis. Agricult.*, vol. 19, no. 4, pp. 735–749, Aug. 2018.

[3] W. Ji, Z. Qian, B. Xu, G. Chen, and D. Zhao, "Apple viscoelastic complex model for bruise damage analysis in constant velocity grasping by gripper," *Comput. Electron. Agricult.*, vol. 162, pp. 907–920, Jul. 2019.

[4] E. Vitzrabin and Y. Edan, "Adaptive thresholding with fusion using a RGBD sensor for red sweet-pepper detection," *Biosyst. Eng.*, vol. 146, pp. 45–56, Jun. 2016.

[5] C. W. Bac, J. Hemming, and E. J. van Henten, "Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper," *Comput. Electron. Agricult.*, vol. 96, pp. 148–162, Aug. 2013.

[6] C. McCool, I. Sa, F. Dayoub, C. Lehnert, T. Perez, and B. Upcroft, "Visual detection of occluded crop: For automated harvesting," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2016, pp. 2506–2513.

[7] S. Kitamura and K. Oka, "A recognition method for sweet pepper fruits using LED light reflections," *Sice J. Control Meas. Syst. Integr.*, vol. 2, no. 4, pp. 255–260, 2009.

[8] C. Zhang, J. Zhang, J. Zhang, and W. Li, "Recognition of green apple in similar background," *Nongye Jixie Xuebao*, vol. 45, no. 10, pp. 277–281, 2014.

[9] S. Sun, Q. Wu, L. Jiao, Y. Long, D. He, and H. Song, "Recognition of green apples based on fuzzy set theory and manifold ranking algorithm," *Optik*, vol. 165, pp. 395–407, Jul. 2018.

[10] Z.-L. He, J.-T. Xiong, R. Lin, X. Zou, L.-Y. Tang, Z.-G. Yang, Z. Liu, and G. Song, "A method of green litchi recognition in natural environment based on improved LDA classifier," *Comput. Electron. Agricult.*, vol. 140, pp. 159–167, Aug. 2017.

[11] W. Ji, X. L. Meng, Y. Tao, B, Xu, and D, Zhao, "Fast segmentation of colour apple image under all-weather natural conditions for vision recognition of picking robots," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 1, pp. 1–9, 2016.

[12] C. W. Bac, J. Hemming, and E. J. van Henten, "Stem localization of sweet-pepper plants using the support wire as a visual cue," *Comput. Electron. Agricult.*, vol. 105, pp. 111–120, Jul. 2014.

[13] X. Liu, D. Zhao, W. Jia, C. Ruan, S. Tang, and T. Shen, "A method of segmenting apples at night based on color and position information," *Comput. Electron. Agricult.*, vol. 122, pp. 118–123, Mar. 2016.

[14] R. Linker, O. Cohen, and A. Naor, "Determination of the number of green apples in RGB images recorded in orchards," *Comput. Electron. Agricult.*, vol. 81, nos. 5–6, pp. 45–57, Feb. 2012.

[15] W. Ji, X. L. Meng, Z. Qian, B. Xu, and D. A. Zhao, "Branch localization method based on the skeleton feature extraction and stereo matching for apple harvesting robot," *Int. J. Adv. Robot. Syst.*, vol. 14, pp. 1–9, May 2017.

[16] C. W. Bac, E. J. van Henten, J. Hemming, and Y. Edan, "Harvesting robots for high-value crops: State-of-the-art review and challenges ahead," *J. Field Robot.*, vol. 31, no. 6, pp. 888–911, 2014.

[17] J. Wei, Q. Zhijie, X. Bo, and Z. Dean, "A nighttime image enhancement method based on Retinex and guided filter for object recognition of apple harvesting robot," *Int. J. Adv. Robot. Syst.*, vol. 15, pp. 1–12, Jan. 2018.

[18] Y. Zhang, Q. Wang, D.-W. Gong, and X.-F. Song, "Nonnegative Laplacian embedding guided subspace learning for unsupervised feature selection," *Pattern Recognit.*, vol. 93, pp. 337–352, Sep. 2019.

[19] T. Yuan, J. Zhang, L. Wei, and Y. Ren, "Feature acquisition of cucumber fruit in unstructured environment using machine vision," *Trans. Chin. Soc. Agricult. Machinery*, vol. 40, no. 8, pp. 170–174, 2009.

[20] J. Lv, F. Wang, L. Xu, Z. Ma, and B. Yang, "A segmentation method of bagged green apple image," *Sci. Hortic*, vol. 246, pp. 411–417, Feb. 2019.

[21] F. DeYao, Y. Qing, Y. BaoJun, Z. YingFeng, G. ZeXin, and T. Jian, "Progress in research on intelligentization of field weed recognition and weed control technology," *Scientia Agricultura Sinica*, vol. 43, no. 9, pp. 1823–1833, 2010.

[22] L. Wang and C. Pan, "Robust level set image segmentation via a local correntropy-based *K*-means clustering," *Pattern Recognit.*, vol. 47, no. 5, pp. 1917–1925, 2014.

[23] A. R. Yadav, R. S. Anand, M. L. Dewal, and S. Gupta, "Multiresolution local binary pattern variants based texture feature extraction techniques for efficient classification of microscopic images of hardwood species," *Appl. Soft Comput.*, vol. 32, pp. 101–112, Jul. 2015.

[24] M. Mehri, P. Héroux, P. Gomez-Krämer, and R. Mullot, "Texture feature benchmarking and evaluation for historical document image analysis," *Int. J. Document Anal. Recognit.*, vol. 20, no. 1, pp. 1–35, 2017.

[25] J. Rakun, D. Stajnko, and D. Zazula, "Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry," *Comput. Electron. Agricult.*, vol. 76, no. 1, pp. 80–88, 2011.

[26] S. Aksoy and R. M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Pattern Recognit. Lett.*, vol. 22, no. 5, pp. 563–582, 2001.

[27] B. Hua, F. L. Ma, and L. C. Jiao, "Research on computation of GLCM of image texture," *Chin. J. Electron.*, vol. 1, no. 1, pp. 155–158, 2006.

[28] Z. Su, B. Tang, Z. Liu, and Y. Qin, "Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine," *Neurocomputing*, vol. 157, pp. 208–222, Jun. 2015.

[29] C.-H. Li, X.-J. Zhu, G.-Y. Cao, S. Sui, and M.-R. Hu, "Identification of the Hammerstein model of a PEMFC stack based on least squares support vector machines," *J. Power Sources*, vol. 175, no. 1, pp. 303–316, 2008.

[30] X. Liu, D. Zhao, W. Jia, W. Ji, and Y. Sun, "A detection method for apple fruits based on color and shape features," *IEEE Access*, vol. 7, pp. 67923–67933, 2019.

[31] S. Ding, L. Xu, C. Su, and F. Jin, "An optimizing method of RBF neural network based on genetic algorithm," *Neural Comput. Appl.*, vol. 21, no. 2, pp. 333–336, 2012.

[32] K. Polat and S. Güneş, "Breast cancer diagnosis using least square support vector machine," *Digit. Signal Process.*, vol. 17, no. 4, pp. 694–701, 2007.

[33] Y. Zhang, D.-W. Gong, and J. Cheng, "Multi-objective particle swarm optimization approach for cost-based feature selection in classification," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 14, no. 1, pp. 64–75, Jan./Feb. 2017.

[34] J. Wei, Z. Jian-Qi, and Z. Xiang, "Face recognition method based on support vector machine and particle swarm optimization," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4390–4393, Apr. 2011.

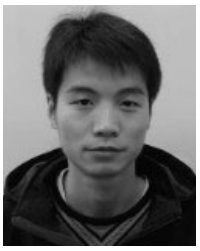
[35] D. Wu, Y. He, S. J. Feng, and Y. D. Bao, "Application of infrared spectra technique based on LS-support vector machines to the non-destructive measurement of fat content in milk powder," *J. Infr. Millim. Waves*, vol. 27, no. 3, pp. 180–184, 2008.

[36] C. Cui, F. Huang, R. Zhang, and B. Li, "Research on cylindricity evaluation based on the particle swarm optimization," *Opt. Precis. Eng.*, vol. 14, no. 2, pp. 256–260, 2006.

[37] C. Singh, E. Walia, and K. P. Kaur, "Enhancing color image retrieval performance with feature fusion and non-linear support vector machine classifier," *Optik*, vol. 158, pp. 127–141, Apr. 2018.


**WEI JI** was born in Henan, China, in 1974. He received the B.Sc. and M.Sc. degrees in electrical engineering from the China University of Mining and Technology, Xuzhou, China, in 1999 and 2002, respectively, and the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 2007.

Since 2007, he has been with the School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, China, where he is currently an Associate Professor. His current research interests include robot for fruit harvesting, machine vision, and artificial intelligence.

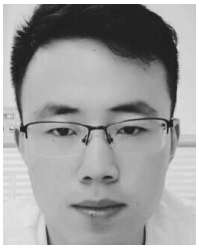


**GUANGYU CHEN** was born in Jiangsu, China, in 1993. He received the B.Sc. degree in automatic control from Jiangsu University, Zhenjiang, China, in 2017, where he is currently pursuing the M.Sc. degree with the School of Electrical and Information Engineering.

His research interests include robotics and machine vision.

**BO XU** was born in Jiangsu, China, in 1977. She received the B.Sc. degree in automatic control from the China University of Mining and Technology, Xuzhou, China, in 1999, the M.Sc. degree in computer science from Jiangsu University, Zhenjiang, China, in 2005, and the Ph.D. degree in electrical engineering from Jiangsu University, Zhenjiang, China, in 2012.

Since 1999, she has been with the School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, China, where she is currently an Associate Professor. Her current research interests include motion control and machine vision.



**XIANGLI MENG** was born in Hebei, China, in 1992. He received the B.Sc. and M.Sc. degrees in automatic control from Jiangsu University, Zhenjiang, China, in 2015 and 2018, respectively.

Since 2018, he has been with the School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, China. His research interests include robotics and machine vision.

**DEAN ZHAO** was born in Jiangsu, China, in 1956. He received the B.Sc. and M.Sc. degrees in electrical engineering from Jiangsu University, Zhenjiang, China, in 1978 and 1982, respectively, and the Ph.D. degree in power electronics from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2006.

Since 1982, he has been with the School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, China, where he is currently a Professor. His research interests include robotics, computers and electronics to agriculture, and in developing harvesting machinery for fruits and vegetables.