

Research Article

Investigation on Recognition Performance of Harvesting Robot Using Regions of Interest Histogram of Oriented Gradients Feature Based on Improved Fuzzy Least Square Support Vector Machine

Jianping Ou  and Jun Zhang 

College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China

Correspondence should be addressed to Jianping Ou; oujianping2018@126.com

Received 28 December 2020; Revised 2 September 2021; Accepted 14 September 2021; Published 6 October 2021

Academic Editor: Ana C. Teodoro

Copyright © 2021 Jianping Ou and Jun Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to solve the problems such as big errors, lack of universality, and too much time consuming occurred in the recognition of overlapped fruits, an improved fuzzy least square support vector machine (FLS-SVM) is established based on the fruit ROI-HOG feature. First, the RGB image is transformed into saturation and value (HSV) image, and then the regions of interest (ROI) are detected from HSV color information. Finally, the histogram of oriented gradients (HOG) feature of ROI will be used as the input of FLS-SVM pattern recognizer to realize the recognition of picking fruit. In addition, the verified FLS-SVM is used to investigate the recognition performance of harvesting robot using regions of interest histogram of oriented gradients feature. The results reveal that the vector sizes are effectively reduced and a higher detection speed is achieved without compromising accuracy relative to conventional approaches. Similarly, the detection accuracy for the learning samples, the isolated fruit, the overlapped fruit, and the background can achieve 99.50%, 96.0%, 89.9%, and 97.0%, respectively, which shows the good performance of the proposed improved ROI-HOG feature recognition method.

1. Introduction

In the visual system of a fruit-picking robot [1, 2], the identification of fruits (including the recognition and orientation of fruits) was of great importance [3–5]. Whether the robot can recognize fruits quickly and precisely or not will have direct influence on its liability and instantaneity.

Some researchers confirmed that it was possible for making recognition and orientation of fruits through the binocular stereo vision. For example, Wei et al. [6] established an automatic extraction method of fruit object under complex agricultural background for vision system of a fruit-picking robot. A lot of experiments show that the extraction accuracy is more than 95%. In addition, Chiu et al. [7] studied an autonomous fruit-picking robot system in greenhouses, and the experimental results showed that the success rates of the integrated picking were in range of

89.63–94.83%. Xiang et al. [8] used binocular stereo vision to recognize clustered tomatoes, and a correct detection rate of 92.4% was achieved. Yang et al. [9] used a convolutional neural network to perform integrated detection of citrus fruits and branches, and the average precisions of fruit and branch recognition were 88.15% and 96.27%, respectively. Tao and Zhou [10] used the fusion of color and 3D feature method to investigate automatic apple recognition for robotic fruit picking, and the result showed that the proposed method exhibited better performance. In addition, Linker et al. [11] determined the number of green apples in RGB images recorded in orchards, and the correct detection rate was more than 95%.

Other researchers [12–14] acquired some certain progress in raising different solutions to proper color space, the fruit image segmentation, feature extraction and orientation, etc. However, they merely analyzed and dealt with the

isolated ones, while there were still limitations to identify whether the fruits overlap each other. Those limitations were expressed as follows. First, the available information which exists in images was not enough; generally, only the information of the grey level, texture, and shape can be employed. Besides, because of their similarity, fruits were not easy to be segmented from each other by existing information. In this situation, related algorithms were desired to find the segmentation line or to rebuild the fruit shape. However, those algorithms were often unique and lack of universality. According to the existed references, the research studies on the identification of fruits mainly concentrated in situations that fruits separate from each other completely, whereas research studies on such more complicated cases were seldom found as fruits were very close or overlapped by others. Yamamoto et al. [15] studied the identification of strawberries that are firmly attached or overlapped by each other. However, this method reduces the errors and improves the accuracy of target identification compared with the traditional method. Subsequently, Anuradha and Sankaranarayanan [16] had proposed a novel mathematical morphological watershed algorithm to preserve these edge details as well as prominent ones to identify tumors in dental radiographs. This algorithm could seek out the boundary to divide tomatoes stacked together and segmented images of ripe tomatoes planted in the field in different growing status automatically. However, the watershed algorithm could fail to mark, so it had a low liability. Xu et al. [17] tried to propose a novel closed-form solution to complete line-segment extraction. The method was tested both on simulated and real-world images, and the results showed that the proposed closed-form solution was feasible in the presence of quantization errors or image noise.

Support vector machine (SVM) has been widely used in many applications [18]. Its goal is to find the optimal separating hyperplane with maximum distance from the closest training samples of each class [19]. Several versions of SVM had been proposed due to its popularity, among which the most important ones were the least squares SVM (LS-SVM). Similarly, among the existing feature extraction methods, the fuzzy least square support vector machine (FLS-SVM) is considered a feasible way and has been extensively used in many applications such as image classification, text classification, and bioinformatics [20]. The method combines the idea of twin SVM and LS-SVM and finds two nonparallel hyperplanes by solving two systems of linear equations instead of nonlinear ones [21]. Although the algorithm provides high accuracies in some applications, both LST-SVM and its improved versions suffer from two main drawbacks: (1) the implicit assumption that sample labels are deterministic, while in many real-world applications, labels come naturally with uncertainties in the form of membership degrees [22]; (2) considering all training samples to be equally important for learning hyperplanes, while in many classification tasks, samples might have different importances [23]. It happens a lot in applications, such as bioinformatics, unbiased class labels, and so on, for example, the task of predicting unknown protein-protein interactions based on computational methods. Large fractions of false

positive interactions have been generated by high throughput methods, which make many of the existing data sets incomplete, incorrect, and noisy [24]. As for a learning machine, weight of the positive interactions generated by the reliable methods is bigger than that of the noisy interactions and outliers [25]. However, the LST-SVM with fuzzy theory is considered a potential approach to cope with these challenges. Fuzzy theory provides useful tools when analyzing complex processes using standard quantitative methods or when the available information is interpreted uncertainly [26]. A fuzzy function offers an efficient way of capturing the inexact nature of real-world problems by representing uncertainty in the data using fuzzy parameters [27].

As mentioned above, the histogram of oriented gradients (HOG) [28–30] based on the region of interest (ROI) [31–33] has a good detection performance. This method requires less eigenvector dimension, so it reduces the calculation amount for training and classifying of feature extraction and classifier, which increases the system speed immensely. However, the FLS-SVM method can improve the two problems of the identification of overlapped fruits picked by a picking robot, including the big errors and long-time processing. The method takes the eigenvector of HOG of the ROI as input for the mode discriminator of the fuzzy least square support vector machine (FLS-SVM). The FLS-SVM has two key features: fuzzifying the parameters and assigning fuzzy membership values. The FLS-SVM replaces the inequality constraints of SVM with equality constraints and converts the quadratic programming into linear system of equations that can be solved by the least square method [34, 35]. All of these have decreased the computing complication and increased the solving speed.

In this paper, a novel FLS-SVM is proposed in order to enhance the detection accuracy of fruit-picking robot. The calculated HOG temperature using four regions of interest (ROI) is used to establish an improved FLS-SVM classifier after the images are detected from HSV color information, which will be useful for recognizing fruits quickly and precisely.

2. An Improved FLS-SVM Recognition Algorithm of Harvesting Robot

The algorithm proposed in this work can identify isolated or overlapped fruits with precision quickly. The flowchart of the improved FLS-SVM recognition algorithm is shown in Figure 1. The identification process mainly consists of the following steps: (1) the images obtained by the camera are converted into the hue, saturation, and value (HSV) color space [36] to extract the ROI domain of fruits; (2) the HOG feature of the ROI field is obtained by the analysis; (3) the obtained HOG feature of the ROI field is input in the mode identifier of FLS-SVM to complete the identification. Moreover, the images are photographed by using the CMOS camera under the sunlight condition. The image processing and FLS-SVM mode identification are completed in the Open Computer Vision Library (OpenCV) and MATLAB R2012a, respectively.

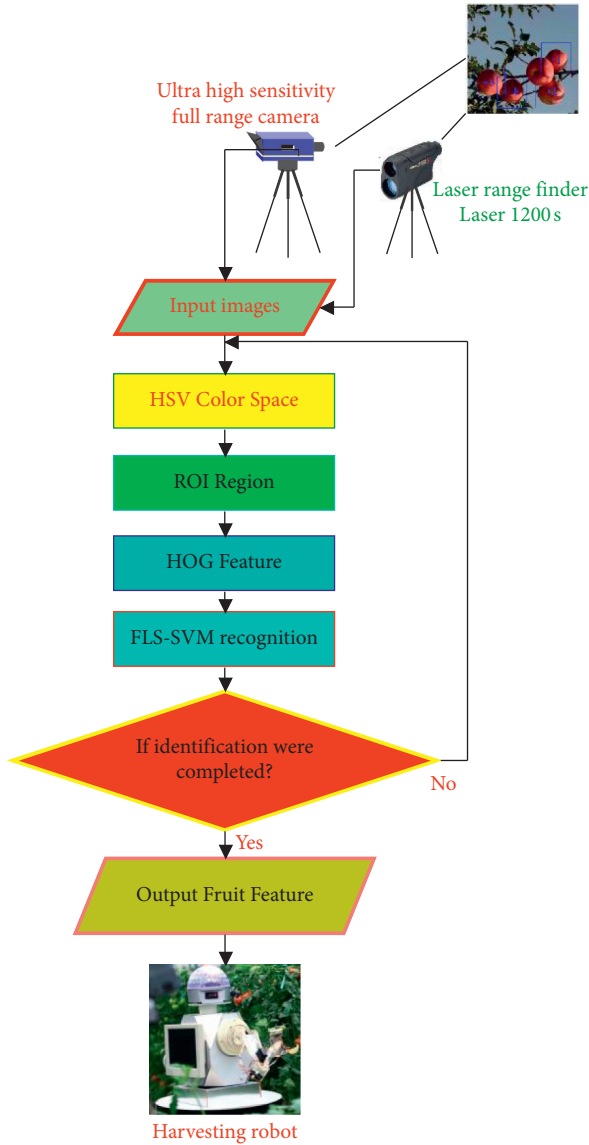


FIGURE 1: Flowchart of the improved FLS-SVM recognition algorithm.

2.1. Conversion of HSV Color Space. In order to ensure the good robustness of color conversion (such as image brightness and contrast), the color conversion cannot be performed directly in the RGB color space. The HSV space, which is the nonlinear transformation of RGB space, converts the values of R, G, and B with strong correlation into values of H, S, and V with weak correlation. The values of H and S are consistent with people’s color sensitivity. There is a certain accordant hue for each uniform color area of the colorful images in the HSV space, which will make the hue available to segment the colorful area solely. The HSV conversion is easy to compute and is also reversible. Besides, the HSV space can meet the attribute requirements of homogeneity, compactness, integrity, naturality, etc. In this paper, the image color space is switched to the HSV by the Open Computer Vision.

The RGB is normalized, and $RGB \in [0, 1]$. The converted hue is also normalized, and $HSV \in [0, 1]$. The transformation formula from RGB to HSV can be expressed as follows:

$$V = \max(R, G, B),$$

$$\delta = V - \min(R, G, B),$$

$$S = \begin{cases} 0, & \text{if } V = 0, \\ \frac{V - \min(R, G, B)}{V}, & \text{else.} \end{cases} \quad (1)$$

Let $r = V - R/\delta$, $g = V - G/\delta$, and $b = V - B/\delta$:

$$6H = \begin{cases} \text{NaN (undefined),} & \text{if } S = 0 \text{ that is } V = 0, \\ 1 - g, & \text{if } V = R, \min(R, G, B) = B, \\ 5 + b, & \text{if } V = R, \min(R, G, B) = G, \\ 3 - b, & \text{if } V = G, \min(R, G, B) = R, \\ 1 + r, & \text{if } V = G, \min(R, G, B) = B, \\ 5 - r, & \text{if } V = B, \min(R, G, B) = G, \\ 3 + g, & \text{if } V = B, \min(R, G, B) = R. \end{cases} \quad (2)$$

When the variation of hue is from 0 to 1.0, the corresponding color is also changed from red to yellow, green, blue-green, blue, amaranth, and finally back to red. More specifically, the values of 0 and 1.0 also represent the red. When the variation of saturation is from 0 to 1.0, the corresponding hue changed from unsaturation (grey shadow) to full saturation (without white content). Similarly, when the variation of brightness is from 0 to 1.0, the corresponding color becomes brighter. Moreover, the value of the brightness is 0 on the conical top. When all the colors are black, the value of saturation is 0, and the value H has no significance by this time.

2.2. Construction of ROI Area. Meanwhile, because of the limitations, the traditional method will certainly increase design calculation workload and big errors and extend design cycle and others, no matter for the extraction of eigenvalue or the training and classification of the classifier if it is used to compute the HOG for the whole color space. For example, when the size of cell is 8×8 , the block size is 16×16 , the block step is 8 pixels, the nine sections are distributed on the gradient direction, and the eigenvector dimensions will be as many as 3780 for the whole color space with a size of 64×128 . In general, the HOG of the fruit central area has little effect on the classification of fruit, and the feature of isolated and overlapped fruit is actually described by the perimeter and area of the fruit [32]. In addition, the HOG in the background region of samples may disturb the classification. In order to solve the problem, an improved method is proposed. When the color space of ROI is converted into HOG, the improved method can improve the recognition rate and reduce computing time. In order to evaluate the segmentation of image objectively, the method

of Liu Jianqing is used [37]. The normalized index can be calculated by the following equation:

$$F(I) = \frac{1}{100(M \times N)} \sqrt{R} \times \sum_{i=1}^R \frac{e_i^2}{\sqrt{A_i}}, \quad (3)$$

where I is the image to be segmented, $M \times N$ is the image size, R is the number of divided regions, and e_i is the color error of the i th region. The smaller the $F(I)$, the better the segmentation effect. In this work, according to the external contour of an apple and method of Liu Jianqing, 5 ROIs (including ROI1, ROI2, ROI3, ROI4, and ROI5) are built, which have basically covered the contour of the apple, as shown in Figure 2.

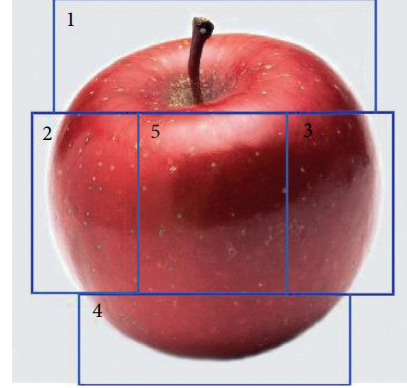


FIGURE 2: Schematic diagram of partition of the apple ROI.

2.3. Extraction of HOG Feature. In general, the feature of HOG is mainly employed to describe the characteristic of local gradient distribution. The good performance can be acquired when it is used to target detection. Thus, the HOG feature can be well used to indicate the characteristic of the local gradient distribution by extracting it. There are 3 steps to describe the extraction of the HOG as shown in Figure 3.

As shown in Figure 3, the pixel of the image is 40×50 , the pixel of the cell is 5×5 , and the block is made up of 2×2 cells. In addition, the value of bin is equal to 9. Therefore, the vector dimension of the HOG is $2268 = 9 \text{ bins} \times (2 \times 2) \text{ cells} \times (7 \times 9) \text{ blocks}$.

- (1) Calculation of amplitude and direction of the gradient:

After the image is grayed, the brightness gradient of X and Y directions is obtained by the following equations:

$$G_x = m(x+1, y) - I(x-1, y), \quad (4)$$

$$G_y = m(x, y+1) - I(x, y-1), \quad (5)$$

where G_x and G_y represent the gradient on the horizontal and vertical direction, respectively; the gradient amplitude $m(x, y)$ reflects the variation of size of the grey level.

In addition, equations (6) and (7) can be used to calculate the amplitude and direction of the gradient, respectively:

$$m(x, y) = \sqrt{G_x^2 + G_y^2}, \quad (6)$$

$$\theta(x, y) = \arctan\left(\frac{G_y}{G_x}\right), \quad (7)$$

where $\theta(x, y)$ represents the gradient direction which can be defined as $0 - \pi$ or $0 - 2\pi$; the gradient direction reflects the variation in direction of the grey level around a certain pixel dot.

- (2) Building of histogram of cell:

The image is divided into $a \times b$ adjacent cells uniformly in its spatial position. Similarly, the HOG has

been optimized on the defined direction within each cell area before it comes to select by using the amplitude or its square or square root of the gradient.

- (3) Description of the block:

In order to explain the variations in illumination and contrast, it is necessary to normalize the gradient amplitude locally. If a block is made up of two adjacent cells, it means that each cell has more than once contributed to the final feature descriptor. The HOG feature descriptor is a vector composed of all the normalized cells within one block.

- (4) Gradient normalization of the block:

Because of the effect of illumination, the variation of gradient amplitude is large, which makes it difficult for the classifier to adapt to its variation. In order to improve the accuracy, the density of each histogram in each block is firstly calculated after obtaining the hog feature vector. Then, each cell in the block is normalized according to the density value. Thus, the normalization of L1-norm and L2-norm can be expressed as follows:

$$v^* = \frac{v}{\sqrt{\|v\|_k + 1.1\epsilon}}, \quad (8)$$

where v is the eigenvector before the normalization, v^* is the eigenvector after the normalization, $\|v\|_k$ represents k -norm, and ϵ is a constant to avoid division by 0.

After the L1-norm and L2-norm are normalized, one more normalization is done for the L1-norm and L2-norm on the ground that the maximum value of the v is set at its threshold such as 0.2.

2.4. Identification Principle of the FLS-SVM. Nowadays, the standard SVM had been extended to the FLS-SVM [38–40], which replaced the inequality constraints of SVM with equality constraints and fuzzy theory. In the samples, n samples are supposed and expressed as $(x_1, y_1, \mu(x_1))$, $(x_2, y_2, \mu(x_2))$, ..., $(x_n, y_n, \mu(x_n))$, $i = 1, 2, \dots, n$, and the constraint condition for the formula is expressed as follows:

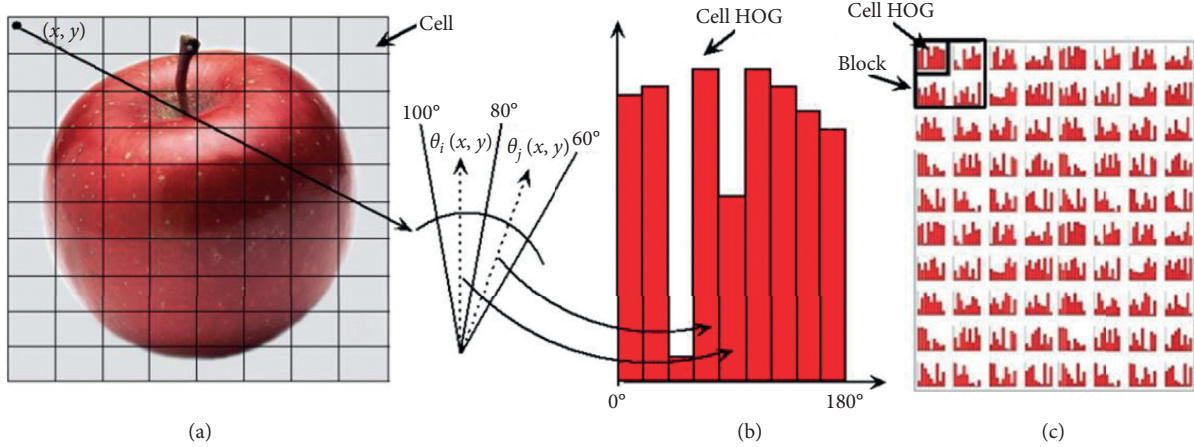


FIGURE 3: Schematic diagram of the extraction of HOG eigenvector. (a) The pixel of the image. (b) The pixel of the cell. (c) The block.

$$\begin{cases} \min & J(\mathbf{w}, \xi_i) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\theta}{2} \sum_{i=1}^n \mu(x_i) \xi_i^2, \\ \text{s.t.} & y_i [\phi(x_i) \times \mathbf{w}^T + b] = 1 - \xi_i, \quad i = 1, 2, \dots, n, \end{cases} \quad (9)$$

where $\xi_i \geq 0$ is the relaxing factor, $\theta > 0$ is the penalty function, and $\mu(x_i)$ is the degree of membership of the x_i , $0 < \mu(x_i) \leq 1$.

Finally, the optimization of FLS-SVM can be transformed to the solution of the linear equations by the method of Lagrange multiplier:

$$L = J - \sum_{i=1}^n a_i [\mathbf{w}^T \phi(x_i) + \xi_i + b - y_i]. \quad (10)$$

In the work, a_i , $i = 1, \dots, n$, and b can be obtained by solving the linear equations. In addition, the classification hyperplane is expressed as follows:

$$\sum a_i y_i K(x, x_i) + b = 0, \quad (11)$$

where $x = x_1, x_2, \dots, x_n$ and $K(\mathbf{x}, x_i) = \exp\{-|\mathbf{x} - x_i|^2 / \sigma^2\}$.

The decision function is finally obtained and expressed as follows:

$$f(x) = \begin{cases} \sum a_i y_i K(x, x_i) + b > 0 & y_1 = 1, \quad \text{apple,} \\ \sum a_i y_i K(x, x_i) + b < 0 & y_1 = -1, \quad \text{environmental background.} \end{cases} \quad (12)$$

The schematic diagram of FLS-SVM identification principle is shown in Figure 4.

3. Results and Discussion

In the section, the improved FLS-SVM is employed to investigate the recognition performance of harvesting robot using regions of interest histogram of oriented gradients feature.

3.1. Optimization of Operating Parameters of the Algorithm

3.1.1. Optimization of HOG Character Parameters. Minetto et al. [41] had developed an effective gradient-based descriptor model for single line text regions and verified the HOG character parameters, including bin, cell, image size, image type, and standardizing types, which have a great influence on the accuracy of identification. In this work, the HOG character parameters are optimized by changing one character parameter while other character parameters are constant. Also, the variations of the HOG character parameters are shown in Table 1.

In order to obtain the optimized image types, the following steps should be carried out. First, other parameters are set as follows: bin = 6, cell = 6 × 6, block = 2 × 2, image size = 90 × 90, and normalization = L1-norm. The HSV image can get a high identification based on the parameter settings above. Then, the other character parameters are optimized. The final results are as follows: bin = 6, cell = 4 × 4, block = 2 × 2, image size = 60 × 60, image type = HSV, and normalization = L1-norm. Finally, according to the optimized parameters of the image sizes, the corresponding ROI area size can be obtained, and the HOG vector size is expressed as shown in Figure 5 when the values of the cell and the block are integers.

3.1.2. Number Optimization of Learning Samples. As the numbers of learning samples increase, the accuracy of identification will increase accordingly. However, computing time may also increase at the same time. Thus, it is very important to find out the proper number of learning samples. In the paper, 10000 figures have been loaded in the experiment in order to evaluate the influence of the number of learning samples on the accuracy of fruit recognition. The influence of the number of learning sample on the identification accuracy is shown in Figure 6. It can be seen that the accuracy for identifying ROI1, ROI2, ROI3, ROI4, and ROI5 reaches 90% when the number of learning samples is up to 4000, which can meet the experimental requirement. In order to improve the computational efficiency and save time, 6000 figures can meet the

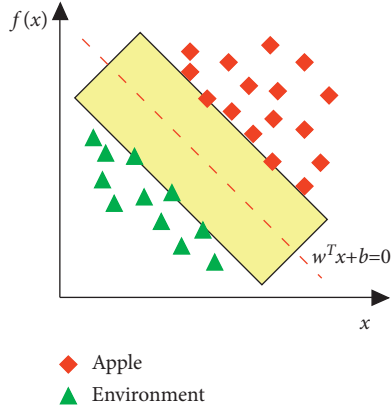


FIGURE 4: Schematic diagram of FLS-SVM identification principle.

TABLE 1: Variation ways of the HOG character parameters.

| Bin | Cell | Block | Image size | Image type | Normalization |
|-----|-------|-------|------------|-------------|---------------|
| 3 | 3 × 3 | 1 × 1 | 90 × 90 | Gray | L1-norm |
| 4 | 4 × 4 | 2 × 2 | 80 × 80 | HSV | L2-norm |
| 5 | 5 × 5 | 3 × 3 | 70 × 70 | RGBmax | L1-Hys |
| 6 | 6 × 6 | 4 × 4 | 60 × 60 | RGB-average | L2-Hys |

experimental requirement and are employed to investigate the optimization of character parameters.

3.1.3. FLS-SVM Parameter Optimization. In the identification model, the selections of penalty factor c and nuclear parameter g affect directly the precision of the model. With the advantage of its objective optimization without using the analyticity of objective function, the self-adaptive genetic algorithm [42] was used to optimize the two presupposed parameters by the prediction model. It is very important to determine the fitness function of adaptive genetic algorithm. The fitness function of adaptive genetic algorithm is expressed as follows:

$$F(c, r) = \frac{1}{\sum_{i=1}^n (y_i - f(x_i))^2 + e}, \quad (13)$$

where y_i is the expected output, $f(x_i)$ is the actual output, and e is a fairly small real number to deal with a possible case in which a denominator becomes zero (less than 0.5×10^{-4}). Selection of crossover probability P_c and mutation probability P_m is critical to the behavior and performance of the genetic algorithm. This paper employed the improved genetic algorithm in which P_c and P_m could change automatically with the adaptive function. The variations of P_c and P_m were expressed as follows:

$$P_c = \begin{cases} 0.95 \times \sqrt{1 - \left(\frac{t}{t_{\max}}\right)^2}, & P_c < 0.5, \\ 0.5, & \text{else,} \end{cases} \quad (14)$$

where t is the genetic algebra and t_{\max} is the terminate algebra.

$$P_m = \begin{cases} 0.01 \times e^{(-\lambda \cdot t / t_{\max})}, & P_m < 0.0005, \\ 0.0005, & \text{else,} \end{cases} \quad (15)$$

where λ is a constant value of 10.

The specific steps for the parameter optimization of FLS-SVM were expressed as follows:

Step 1: choose the learning samples and testing samples. Then, set the intervals of the penalty factor and the nuclear factor as (0, 100) and (0, 10), respectively. Finally, the initial SVM parameter groups are produced.

Step 2: set the values of the crossover probability, mutation probability, size of the group, and generations as 0.6, 0.2, 50, and 100, respectively.

Step 3: conduct the training. The evolutionary curve of the fitness that varies with the process of searching for the best parameter by the genetic algorithm is shown in Figure 7. The final optimized values of the penalty factor c and the nuclear parameter g are as follows: $c = 40.5678$ and $g = 3.5635$.

3.2. Experiment Results. The operating conditions for this experiment were as follows: PC: CPU Intel® Xeon® 2.00 GHz, Memory 4.0 GB; OS: Windows XP, Visual Studio 2010, OpenCV 2.2, and MATLAB 2012a LIBSVM. In order to verify the superiority of this algorithm for fruits identification, the isolated fruits' image and the overlapped fruits' image were placed, respectively, into the learning samples and the testing samples.

In order to show the advantages of using FLS-SVM instead of SVM, the comparisons of experimental identification results between the FLS-SVM, the SVM, and the faster R-CNN [43, 44] are shown in Figure 8 and Table 2. Obviously, the FLS-SVM identification method improved the identification accuracy compared with the SVM identification method. More specifically, the identification accuracies of the FLS-SVM identification method are 99.50%, 96.00%, 89.90%, and 97.00% for the learning samples, the isolated fruits, overlapped fruits, and the environmental background, respectively. However, the identification accuracies of the SVM identification method are 98.60%, 93.10%, 85.60%, and 95.00% for the learning samples, the isolated fruits, overlapped fruits, and the environmental background, respectively. Moreover, the identification accuracies of the faster R-CNN identification method are 99.48%, 95.80%, 89.00%, and 96.87% for the learning samples, the isolated fruits, the overlapped fruits, and the environmental background, respectively.

Identification speed comparison between the FLS-SVM identification method and the fast R-CNN method is shown in Table 3. It can be concluded that identification speed of the FLS-SVM identification method is faster than that of the fast R-CNN method.

As previously mentioned, the identification algorithm using FLS-SVM proposed in this work based on the ROI and HOG had shown a good performance, which is consistent with the results obtained in the literature [45].

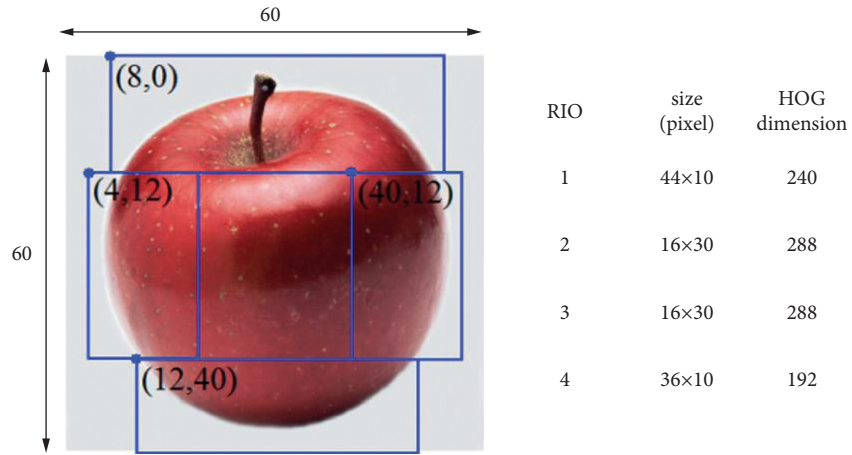


FIGURE 5: ROI area size and the number of HOG vector dimension.

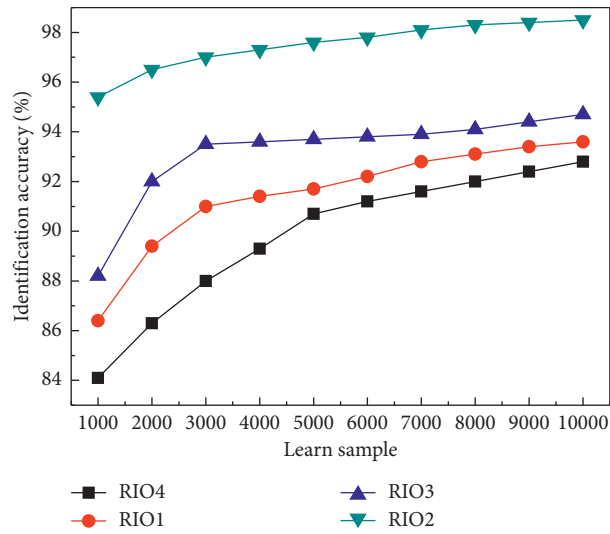


FIGURE 6: Influence of the number of learning sample on the identification accuracy.

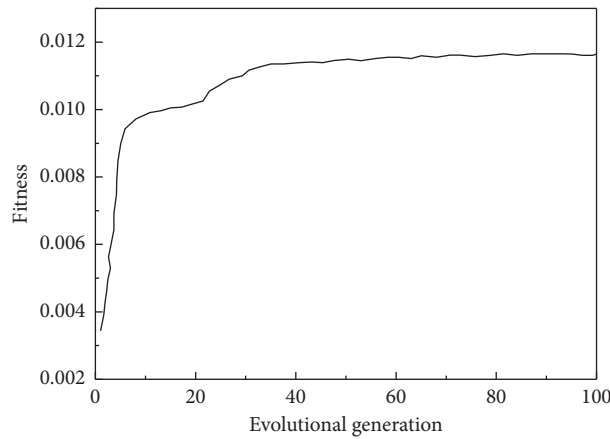


FIGURE 7: Fitness curve varies with the search for the best parameter.

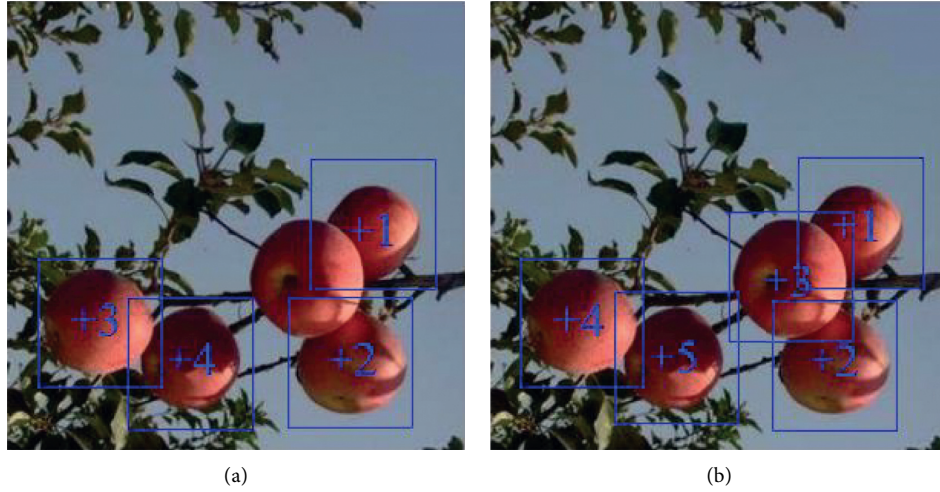


FIGURE 8: Results of the (a) incorrect and (b) correct identification of fruit.

TABLE 2: Comparisons of experimental identification results between the FLS-SVM, the SVM, and the faster R-CNN.

| Identification methods and identification results | Identification images | | | |
|---|-----------------------|----------------------------|----------------------|----------------------------------|
| | Learning samples | Isolated fruit test sample | Overlap fruit sample | Environmental background samples |
| Image total | 4000 | 2000 | 1000 | 3000 |
| Correct amount | 3945 | 1862 | 856 | 2850 |
| SVM identification results | | | | |
| Number of errors for environmental background | 45 | 110 | 110 | 30 |
| Number of errors for fruit background | 10 | 28 | 34 | 120 |
| Accuracy | 98.60% | 93.10% | 85.60% | 95.00% |
| FLS-SVM identification results | | | | |
| Correct amount | 3982 | 1920 | 894 | 2910 |
| Number of errors for environmental background | 16 | 80 | 106 | 0 |
| Number of errors for fruit background | 2 | 0 | 0 | 90 |
| Accuracy | 99.50% | 96.00% | 89.40% | 97.00% |
| Faster R-CNN identification results | | | | |
| Correct amount | 3979 | 1916 | 890 | 2906 |
| Number of errors for environmental background | 17 | 82 | 108 | 2 |
| Number of errors for fruit background | 4 | 2 | 2 | 92 |
| Accuracy | 99.48% | 95.80% | 89.00% | 96.87% |

TABLE 3: Identification speed comparison between the FLS-SVM identification method and the fast R-CNN method.

| Identification items | Identification method | |
|------------------------------------|-----------------------|------------|
| | FLS-SVM | Fast R-CNN |
| Training time | 8.85 h | 9.48 h |
| Identification time for each image | 0.083 s | 0.31 s |

4. Conclusions

In recent years, artificial intelligence technology [46, 47] has been rapidly developed and widely used in all walks of life. In

this work, an algorithm for solving the problem of recognition of overlapped fruits is proposed and evaluated based on fuzzy least squares support vector regression. The RGB image is transformed into HSV image, and then the regions

of interest are detected from HSV color information. In the numerical experiments, the main conclusions are as follows:

- (1) This algorithm is able to effectively reduce the HOG vector dimension and the processing speed by using ROI and can also improve the identification accuracy by combining the HOG feature which was used to describe the local gradient distribution and the FLS-SVM.
- (2) The verified FLS-SVM is used to investigate the recognition performance of harvesting robot using regions of interest histogram of oriented gradients feature. The vector sizes are effectively reduced and a higher detection speed is achieved without compromising accuracy relative to conventional approaches.
- (3) The algorithm is a pragmatic identification method for picking a robot to pick fruits. Compared with the conventional method, the detection accuracy of FLS-SVM identification method for the learning samples, the isolated fruit, the overlapped fruit, and the environmental background can achieve 99.50%, 96.0%, 89.9%, and 97.0%, respectively, which shows the good performance of the proposed improved ROI-HOG feature recognition method.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 61471370.

References

- [1] Y. Xiong, Y. Ge, and P. J. From, "An obstacle separation method for robotic picking of fruits in clusters," *Computers and Electronics in Agriculture*, vol. 175, Article ID 105397, 2020.
- [2] Z. Li, F. Miao, Z. Yang, and H. Wang, "An anthropometric study for the anthropomorphic design of tomato-harvesting robots," *Computers and Electronics in Agriculture*, vol. 163, Article ID 104881, 2019.
- [3] Y. Miao and J. Zheng, "Optimization design of compliant constant-force mechanism for apple picking actuator," *Computers and Electronics in Agriculture*, vol. 170, Article ID 105232, 2020.
- [4] C. Liang, J. Xiong, Z. Zheng et al., "A visual detection method for nighttime litchi fruits and fruiting stems," *Computers and Electronics in Agriculture*, vol. 169, Article ID 105192, 2020.
- [5] F. Gao, L. Fu, X. Zhang et al., "Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN," *Computers and Electronics in Agriculture*, vol. 176, Article ID 105634, 2020.
- [6] X. Wei, K. Jia, J. Lan, Y. Li, Y. Zeng, and C. Wang, "Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot," *Optik*, vol. 125, no. 19, pp. 5684–5689, 2014.
- [7] Y.-C. Chiu, S. Chen, and J.-F. Lin, "Study of an autonomous fruit picking robot system in greenhouses," *Engineering in Agriculture, Environment and Food*, vol. 6, no. 3, pp. 92–98, 2013.
- [8] R. Xiang, H. Jiang, and Y. Ying, "Recognition of clustered tomatoes based on binocular stereo vision," *Computers and Electronics in Agriculture*, vol. 106, pp. 75–90, 2014.
- [9] C. H. Yang, L. Y. Xiong, Z. Wang et al., "Integrated detection of citrus fruits and branches using a convolutional neural network," *Computers and Electronics in Agriculture*, vol. 174, Article ID 105469, 2020.
- [10] Y. Tao and J. Zhou, "Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking," *Computers and Electronics in Agriculture*, vol. 142, pp. 388–396, 2017.
- [11] R. Linker, O. Cohen, and A. Naor, "Determination of the number of green apples in RGB images recorded in orchards," *Computers and Electronics in Agriculture*, vol. 81, pp. 45–57, 2012.
- [12] P. Zhao and N.-H. Wang, "Precise perimeter measurement for 3D object with a binocular stereo vision measurement system," *Optik*, vol. 121, no. 10, pp. 953–957, 2010.
- [13] P. Zhao and G. Q. Ni, "Simultaneous perimeter measurement for 3D object with a binocular stereo vision measurement system," *Optics and Lasers in Engineering*, vol. 48, no. 4, pp. 505–511, 2010.
- [14] Y.-Z. Gu, M. Sato, and X.-L. Zhang, "An active stereo vision system based on neural pathways of human binocular motor system," *Journal of Bionics Engineering*, vol. 4, no. 4, pp. 185–192, 2007.
- [15] K. Yamamoto, S. Ninomiya, Y. Kimura, A. Hashimoto, Y. Yoshioka, and T. Kameoka, "Strawberry cultivar identification and quality evaluation on the basis of multiple fruit appearance features," *Computers and Electronics in Agriculture*, vol. 110, pp. 233–240, 2015.
- [16] K. Anuradha and K. Sankaranarayanan, "Detection of oral tumor based on marker controlled watershed algorithm," *International Journal of Computer Applications*, vol. 52, no. 2, pp. 15–18, 2012.
- [17] Z. Xu, B.-S. Shin, and R. Klette, "Closed form line-segment extraction using the Hough transform," *Pattern Recognition*, vol. 48, no. 12, pp. 4012–4023, 2015.
- [18] J. S. Sartakhti, H. Afrabandpey, and N. Ghadiri, "Fuzzy least squares twin support vector machines," *Engineering Applications of Artificial Intelligence*, vol. 85, pp. 402–409, 2019.
- [19] S. Sun, X. Xie, and C. Dong, "Multiview learning with generalized eigenvalue proximal support vector machines," *IEEE Transactions on Cybernetics*, vol. 99, pp. 1–10, 2018.
- [20] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [21] Z. Liu, C. Chen, and Y. Zhang, "Interval type-2 fuzzy weighted support vector machine learning for energy efficient biped walking," *Applied Intelligence*, vol. 40, no. 3, pp. 453–463, 2014.
- [22] M. Tanveer, M. A. Khan, and S.-S. Ho, "Robust energy-based least squares twin support vector machines," *Applied Intelligence*, vol. 45, no. 1, pp. 174–186, 2016.

- [23] Z. Zhang, T. Lv, H. Wang, L. Liu, and J. Tan, "A novel least square twin support vector regression," *Neural Processing Letters*, vol. 48, pp. 1–14, 2018.
- [24] X. Xie, "Regularized multi-view least squares twin support vector machines," *Applied Intelligence*, vol. 48, no. 9, pp. 3108–3115, 2018.
- [25] X. Xie, S. Sun, H. Chen, and J. Qian, "Domain adaptation with twin support vector machines," *Neural Processing Letters*, vol. 48, no. 2, pp. 1213–1226, 2018.
- [26] X. Zhang and X. Gao, "Twin support vector machines and subspace learning methods for microcalcification clusters detection," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 5, pp. 1062–1072, 2012.
- [27] E. D. Übeyli, "Support vector machines for detection of electrocardiographic changes in partial epileptic patients," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1196–1203, 2008.
- [28] V. Alilou and F. Yaghmaee, "Non-texture image inpainting using histogram of oriented gradients," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 43–53, 2017.
- [29] A. Kumar, "Deblurring of motion blurred images using histogram of oriented gradients and geometric moments," *Signal Processing: Image Communication*, vol. 55, pp. 55–65, 2017.
- [30] R. Kapoor, R. Gupta, L. H. Son, S. Jha, and R. Kumar, "Detection of power quality event using histogram of oriented gradients and support vector machine," *Measurement*, vol. 120, pp. 52–75, 2018.
- [31] N. Kouadria, K. Mechouek, S. Harize, and N. Doghmane, "Region-of-interest based image compression using the discrete Tchebichef transform in wireless visual sensor networks," *Computers & Electrical Engineering*, vol. 73, pp. 194–208, 2019.
- [32] Z. Zuo, X. Lan, L. Deng, S. Yao, and X. Wang, "An improved medical image compression technique with lossless region of interest," *Optik*, vol. 126, no. 21, pp. 2825–2831, 2015.
- [33] E. Enjilela and E. M. A. Hussein, "Refining a region-of-interest within an available CT image," *Applied Radiation and Isotopes*, vol. 75, pp. 77–84, 2013.
- [34] T.-S. Wang and H.-Y. Zuo, "Fuzzy least squares support vector machine soft measurement model based on adaptive mutative scale chaos immune algorithm," *Journal of Central South University*, vol. 21, no. 2, pp. 593–599, 2014.
- [35] Y. Wang, X. Zhao, and B. Wang, "LS-SVM and Monte Carlo methods based reliability analysis for settlement of soft clayey foundation," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 5, no. 4, pp. 312–317, 2013.
- [36] Y. Xu, K. Imou, Y. Kaizu, and K. Saga, "Two-stage approach for detecting slightly overlapping strawberries using HOG descriptor," *Biosystems Engineering*, vol. 115, no. 2, pp. 144–153, 2013.
- [37] J. Jianqing Liu and Y. H. Yee-Hong Yang, "Multiresolution color image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 689–700, 1994.
- [38] H.-Y. Zuo, Z.-Q. Luo, J.-L. Guan, and Y.-W. Wang, "Identification on rock and soil parameters for vibration drilling rock in metal mine based on fuzzy least square support vector machine," *Journal of Central South University*, vol. 21, no. 3, pp. 1085–1090, 2014.
- [39] C. Qian, H. Zhu, Q. Peng, W. Zuo, and G. Liu, "Parameters identification investigations on the hysteretic Preisach model improved by the fuzzy least square support vector machine based on adaptive variable chaos immune algorithm," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 36, no. 3, pp. 227–242, 2017.
- [40] S. Zhang, W. Lu, W. Xing, and L. Zhang, "Using fuzzy least squares support vector machine with metric learning for object tracking," *Pattern Recognition*, vol. 84, pp. 112–125, 2018.
- [41] R. Minetto, N. Thome, M. Cord, N. J. Leite, and J. Stolfi, "T-HOG: an effective gradient-based descriptor for single line text regions," *Pattern Recognition*, vol. 46, no. 3, pp. 1078–1090, 2013.
- [42] H. Yang and F. Deng, "A novel adaptive mutative scale optimization algorithm based on chaos genetic method and its optimization efficiency evaluation," *Journal of Central South University of Technology*, vol. 19, no. 9, pp. 2554–2560, 2012.
- [43] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [44] P. Agrawal, R. Girshick, and J. Malik, "Analyzing the performance of multilayer neural networks for object recognition," in *Proceedings of the Computer Vision-ECCV 2014*, pp. 329–344, Zurich, Switzerland, September 2014.
- [45] Z. J. Fan and X. X. Li, "Recognition of potato diseases based on fast detection and fusion features of ROI," *Southwest China Journal of Agricultural Sciences*, vol. 32, no. 3, pp. 544–550, 2019.
- [46] H. Zuo, B. Zhang, Z. Huang, K. Wei, and J. Tan, "Effect analysis on SOC values of the power lithium manganate battery during discharging process and its intelligent estimation," *Energy*, vol. 238, Article ID 121854, 2022.
- [47] E. Jiaqiang, B. Zhang, Y. Zeng et al., "Effects analysis on active equalization control of lithium-ion batteries based on intelligent estimation of the state-of-charge," *Energy*, vol. 238, Article ID 121822, 2022.