Fruit recognition based on pulse coupled neural network and genetic Elman algorithm application in apple harvesting robot

Weikuan Jia\textsuperscript{1,2}, Shanhao Mou\textsuperscript{1}, Jing Wang\textsuperscript{1}, Xiaoyang Liu\textsuperscript{2}, Yuanjie Zheng\textsuperscript{1,3}, Jian Lian\textsuperscript{4} and Dean Zhao\textsuperscript{2}

Abstract
In order to improve the harvesting efficiency of apple harvesting robot, this article presents an apple recognition method based on pulse coupled neural network and genetic Elman neural network (GA-Elman). Firstly, we use pulse coupled neural network to segment the captured 150 images, respectively, and extract six color features of R, G, B, H, S, and I and 10 shape features of circular variance, density, the ratio of perimeter square to area, and Hu invariant moments of segmented images, and these 16 features are considered as the inputs of Elman neural network. In order to overcome some defects of Elman neural network, such as, trapping local minimum easily and determining the number of hidden neurons difficultly; in this article, genetic algorithm is introduced to optimize it, and new optimization way is designed, that is, the connection weights and number of hidden neurons separate encoding and evolving simultaneously, in the process of structural evolution at the same time the learning of connection weights is completed, and then the operating efficiency and recognition precision of Elman model are improved. In order to get more precision neural network, and avoid the influence of fruit recognition caused by branches or leaves shadow, apple along with branches and leaves is allowed to train. The results of experiments show that compared with the traditional back-propagation, Elman neural network, and other two recognition algorithms of obscured fruit. the genetic Elman neural network algorithm is the optimal method which successful training rate can reach to 100%, recognition rate of overlapping fruit and obscured fruit can reach to 88.67% and 93.64%, respectively, and the total recognition rate reaches to 94.88%.

Keywords
Apple harvesting robot, PCNN segmentation, GA-Elman neural network, fruit recognition

Date received: 11 July 2019; accepted: 29 November 2019

Corresponding authors:
Weikuan Jia, School of Information Science and Engineering, Shandong Normal University, Daxue road No.1, Jinan City, Shandong Province 250358, China.
Email: jwk_1982@163.com

Yuanjie Zheng, School of Information Science and Engineering, Shandong Normal University, Daxue road No.1, Jinan City, Shandong Province 250358, China.
Email: yjzheng@sdu.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Introduction

With the advancement of science and technology, robotics technology has gradually matured, and apple harvesting robot also makes rapid growth.\(^1,2\) Apple harvesting robot is relatively complex, and the working environment of the visual system as the “eyes” of harvesting robot is a key technology of harvesting robot intelligence. Accompanied by either monocular vision system\(^3\) or binocular vision system,\(^4\) whether identify static fruit,\(^5\) or dynamic fruit,\(^6\) regardless of single fruit,\(^7\) overlapped fruit,\(^8\) and in the near-scene apple,\(^9\) to recognize apple in the night,\(^10\) these have made great progress. Kelman et al. proposed convexity of fruit tree images detection to determine the apple edges and 3-D modeling by the least square constraint mechanism, respectively, fixing the mature individual apple, thereby to reach an accuracy of 94\%.\(^11\) Wachs et al. used the maximum interactive information of two forms of images, infrared and color images, to extract high-level and low-level visual features, and in fruit canopy detection, the two feature recognition rates of “green” apples can reach 54\% and 74\%, respectively.\(^12\) Zhang et al. utilizes support vector machine and super green operator to propose the recognition of nearly color background green apple to achieve the recognition rate of 89.3\% in the night;\(^9\) Li et al. uses binocular stereo vision technology for apple’s three-dimensional position and reduces the search scope and other measures to improve around two image matching precision, with the range error about 5\%.\(^13\) The rapid and efficient recognition of apple fruit affects the operating efficiency of harvesting robots directly. However, the image segmentation and object recognition are the bottlenecks restricting its development, which affect the reliability and real-time performance of harvesting robots.

Image segmentation is the premise of feature extraction and target recognition, and the results have a direct impact on the progress of follow-up work.\(^14,15\) For apple image segmentation, in the past, the segmentation algorithm is based on threshold or improved threshold,\(^16,17\) such as K-means clustering,\(^18\) Otsu,\(^19\) and so on, and these methods have achieved good results. However, they mainly depend on image gray level information but ignored the spatial information of image. In this study, the pulse coupled neural network (PCNN)\(^20\) is introduced into the apple image segmentation; the advantage is that both the spatial information and the gray information of the image are considered to obtain a more ideal segmentation. Accurately and rapidly recognition of apple target is still difficult and the key point of apple harvesting robots vision system. In order to further improve the recognition accuracy and speed of apple harvesting robots and meet real-time requirements of apple harvesting operations, it needs to further optimize recognition algorithms. Elman neural network is introduced into apple image recognition, and this model is a feedback neural network with the ability of adapting to time-varying dynamic characteristic and strong global stability. Genetic algorithm (GA) is widely used in optimization.\(^21,22\) In order to overcome the inherent defects and improve the performance of Elman neural network, GA is introduced, and a new optimization method is adopted, that is, the connection weights and hidden neurons are separate encoding and evolving simultaneously, and thus GA-Elman neural network is obtained.

In order to further improve the working efficiency of apple harvesting robot, in this article, a new target fruit recognition method is proposed. First, PCNN algorithm is employed to segment the apple images, which can take into account the color features and spatial information of the images, and the segmentation effect is ideal. Then, the segmented images are processed by corresponding morphological processing, the 6 color features and 10 shape features were extracted, and these 16 features are considered as the typical features of target apples. Finally, a new GA-Elman algorithm is established and used to recognize the target apples, which can improve the operating efficiency and recognition precision.

The rest of the article is organized as follows: In the second section, firstly, the brief introduction of the apple image acquisition is given, the PCNN segmentation algorithm is described, target fruit segmented image is obtained, and then some morphological processing is applied in order to extract the color and shape features. In the third section, the new thought of GA-Elman algorithm is described, and then to construct a new simultaneous optimization way of connection weight and network structure, the basic steps of new classification algorithm are given. In the fourth section, the experiment is arranged, and the performance of new method is evaluated from multiple perspectives, such as, recognition rate, training times, recognition times, and recognition rate of fruit with different postures. The fifth section provides some conclusion and discussion of the new method.

Apple images segmented by PCNN

In order to better obtain the characteristic information of the apple image and provide the basis for the identification of the target fruit, there is a need to do a series of processing on the image, such as image denoising, image segmentation, feature extraction, and so on. The image segmentation and feature extraction are directly related to the recognition accuracy of the object.

Image acquisition

Image acquisition location: Images were acquired from apple demonstration bases in Dashaihe Town, Fengxian County, Jiangsu Province, Jiangsu University apple harvesting robot experiment bases.

Image acquisition information: all images are collected under natural light, including sunny and cloudy conditions, and smooth and backlighting, all the images are collected between 8:00 and 17:00 pm.
Image acquisition device: the camera is connected to the computer, AFT-0814MP (Vision Electronics Co., Ltd), with focal length 8 mm, field of view angle 54°, and closest object distance 0.15 m, and it is made sure that the image acquisition distance is between 50 and 200 cm.

The variety of the apples is Fuji, and a total of 150 apple images were collected in the experiment which include 97 sunny and 53 cloudy, and the number of the fruits is 229. The collected color image of the apple is shown in Figure 1.

In the natural environment, the natural growth posture of apple fruit changes, and the position factors of the image acquisition system of the harvesting robot make the collected apple images to vary, such as single unobscured fruit, adjacent fruit, overlapping fruit, and foliage cover the fruit. The light environment of the collected image and the growth posture of the fruit will affect the recognition efficiency and accuracy of the visual system, thus affect the harvesting efficiency.

**PCNN segmentation algorithm**

As a third-generation neural network, PCNN is widely used in image processing. PCNN is a self-organizing network that does not require training and the network was constructed by simulating the activities of the mammal’s visual cortex neurons. Through the interaction of neighboring pixels, the neuron can simultaneously pulse with similar input data, which can compensate for the spatial incoherence, time series feature, and small changes in amplitude of input data, thereby completely retaining the area information of the image.23-26

When using PCNN for image segmentation, the two-dimensional image matrix is regarded as the PCNN neuron models, and the gray value of each pixel corresponds to the input of each neuron. If there are some pixels with similar gray values in the neighborhood of internal connection weight matrix $M$, the pulse output generated by one of the pixels will cause neurons active corresponded by nearby other gray-scale pixels, that is, generated pulse output. The binary image is the output segmentation image.

Therefore, PCNN can group image pixels based on spatial proximity and brightness similarity and apply it to image segmentation, which can achieve better segmentation results. Currently, PCNN image segmentation has achieved good results in many fields. This study is applied to apple image segmentation.

The simplified model of PCNN neurons is shown in Figure 2.

The mathematical description of the simplified PCNN neurons model is as follows:

In the input area

$$F_{ij}(n) = I_{ij}$$

(1)

$F_{ij}$ is the neuron input, and $I_{ij}$ is the external nerve input excitation, that is, pixel $(i, j)$ corresponds to characteristic value $L_{ij}(n)$

$$L_{ij}(n) = \sum W_{ij} Y_{ij}(n-1)$$

(2)

$L_{ij}$ is the connection entry, $Y_{ij}$ is the pulse output, and $W_{ij}$ is the connection coefficient of coupling connection.

In the connection input area

$$U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n))$$

(3)

$U_{ij}$ is the internal active item, and $\beta$ is the connection coefficient for internal active items.

In the pulse generator stage

$$Y_{ij}(n) = \begin{cases} 1 & (U_{ij}(n) > E_{ij}(n-1)) \\ 0 & (U_{ij}(n) \leq E_{ij}(n-1)) \end{cases}$$

(4)

$E_{ij}$ is the dynamic threshold, and the expression of $E_{ij}$ is as follows

$$E_{ij}(n) = E_{ij}(n-1) - \gamma + \nu Y_{ij}(n)$$

(5)

In formula, $\gamma$ is the dynamic threshold attenuation step and $\nu$ is the dynamic threshold amplification factor.
Segmentation effect of apple images

Figure 1 shows an example and the segmented results are shown in Figure 3.

At present, there are no standards for judging the effect of image segmentation, and most of them depend on intuitive visual. From the results of image segmentation, for background where the space distances are closer and the color differences are similar, it is segmented into the same class. Segmented images of target apple, leaves, and branches are clear, with good visual effects. In this study, PCNN algorithm is introduced, which takes into account the spatial location of the pixels and color difference information. The segmentation result is more ideal.

Good segmentation results are beneficial to the feature extraction of each segmented image, and the extracted features are more efficient. If the segmentation effect is poor, it may cause the extracted features not to express the characteristics of the target, or the extracted features are incomplete, and it is even difficult to extract the effective features of the target, which ultimately influence the recognition effect.

Feature extraction of target fruit

Feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and nonredundant, facilitating the subsequent learning and generalization steps and, in some cases, leading to better human interpretations. The training process of the recognition model is based on the training samples consisting of a single fruit and focuses on the completeness of the extracted apple fruit characteristics to fully characterize the apple features. Therefore, we try to choose a single unobscured fruit as a training sample in the selection of training samples. In the process of identification, due to the diversity of apple fruits, the characteristics of the samples extracted from the fruits are not complete, such as blocking and overlapping of fruits and trees, leading to information loss and affecting the identification of fruits. In addition, the lighting environment will also interfere with the recognition. As the shielding area and occluding edge are different, the extracted features are also very different and will affect the fruit recognition. In order to avoid the influence of the above factors, and obtain a more accurate identification neural network model, we train the model with apple fruits, branches and leaves together.

Therefore, feature extraction of training samples must consider not only the apple fruit but also the feature extraction of branches and leaves. We need to intercept a relatively complete leaf and a branch of a branch in each image. When identifying sample image processing, we only extract the characteristics of the target fruit. Depending on the characteristics of apples, leaves, and branches, this study considers the extraction of its color and shape characteristics to better describe the object.

Color feature. The color feature of an image is a global feature used to describe the surface properties of the scene corresponding to the image or image region. It is based on the characteristics of the pixels, and all pixels belonging to the image or image region have their own contributions. The color difference of apple fruit, leaves, and branches is relatively obvious, so the color characteristics of the extracted objects are practical and feasible.

In red-green-blue (RGB) color space, the color can be directly compared, and the characteristic information of the color image object of R, G, and B can be well controlled, and the characteristics of these three components is very effective in the color feature extraction. Because the apple harvesting robot belongs to outdoor operations, we need to consider the influence of light and further extract color features in the hue-saturation-intensity (HSI) color space.
HSI color space is similar to the mechanism of human color perception and can adapt to all lighting conditions. The three component features of the image object, such as H, S, and I, are necessary in color feature extraction. In this study, six component features in the RGB and HSI color spaces were extracted to describe the global features of the target.

Shape feature. Because the color is insensitive to changes in the direction and size of the image, color features cannot express the local features of the image object well, so it is necessary to further extract the shape features. Apple fruit, leaves, and branches have their own specific shapes, and the differences are more obvious. Therefore, the shape characteristics of the extracted objects are practical and feasible.

According to the characteristics of the segmented image, we select image geometry characteristics of the object and the Hu invariant moment, in which the geometric feature mainly selects the circle variability, density, and perimeter square area ratio of three eigenvectors, and Hu invariant moment contains seven feature vectors. In this study, a total of 10 shape features were extracted and their biggest feature is that they do not change because of the position, size, and angle of the target in the image. They satisfy the RST (rotation, proportion, and translation) invariance. Hu invariant moments are put into use for extracting the shape features of each connected area in target image, and the seven invariant moments are as follows.

Figure 3. Apple image PCNN segmentation results. (a) Segmented apple; (b) Split leaves; (c) Split tree branches. I, II, III correspondence to Figure 1(a), 1(b), and 1(c). PCNN: pulse coupled neural network.
Hu invariant moments ensure the rotation, translation, and scaling invariance of fields. However, Elman neural network carries forward Elman neural network has been widely used in many fields. This study uses the Elman neural network design classifier to use the 16 features extracted in the third section as input to the neural network for training modeling.

**GA-Elman recognition algorithm**

The recognition of apple fruit can be regarded as a classification problem in essence, while the classification problem is an input/output transformation. The neural network model has unique advantages in dealing with such problems. This study uses the Elman neural network design classifier to use the 16 features extracted in the third section as input to the neural network for training modeling.

**Thought of GA-Elman algorithm**

Elman neural network has been widely used in many fields. However, Elman neural network carries forward the advantages of the back-propagation (BP) algorithm and, at the same time, it also inevitably inherited the some inherent disadvantages of the BP network, such as easy to trap in local minimum, resulting in training failure; the learning rate is fixed, which limits the network’s convergence rate; the number of hidden neurons is difficult to determine, thus artificial attempts to waste a lot of time. These deficiencies limit the neural network transport efficiency and recognition accuracy. So, the genetic algorithm is introduced into the Elman neural network, and its optimization is performed in two aspects: first, to optimize the connection weights and thresholds of the network, to improve the network learning efficiency, and to solve the fixed learning rate and the shortcomings of the local minimum; second, to optimize the hidden layer neurons numbers, which can solve the problem in determining the difficult hidden layer structure, and to reduce the time people try to construct the network.

In the previous studies with evolutionary Elman neural network of genetic algorithms, most of them optimize the connection weight or network structure (the number of hidden neurons) individually. It has good results in enhancing network performance, but the effect is still slightly unsatisfactory. In the optimization of weights of space, although it obtains the optimal connection weights and thresholds, this needs to be based on the specific network structure; simply optimizing the network structure can save people from trying to determine the number of hidden neurons, but still a weight learning can’t get rid of the shortcomings of the inherited BP algorithm. Therefore, the study proposes a novel optimization method. Elman neural network connection weights and the hidden neurons are coded and simultaneously evolved, constructing a new classification algorithm based on genetic algorithm optimization (GA-Elman) to improve the classification efficiency and accuracy of the model.

**Elman neural network**

Elman neural network is based on back-BP neural network optimization, adding an undertaking layer in the hidden layer of BP network, and as there is a time delay operator to achieve the purpose of memory, the system has a time-varying dynamic characteristics ability and strong global stability. Elman neural network is a feedback neural network, and its topology structure is shown in Figure 4.

The topology structure of Elman neural network is generally made up of four layers: input layer, hidden layer, undertaking layer, and output layer. The undertaking layer is used to memorize the output value of the hidden layer unit at the previous moment and can be considered as one-step delay operator. Based on the BP network structure, the output of the hidden layer is delayed and stored by the receiving layer and self-linked to the input of the hidden layer. This self-association-mode makes it sensitive to historical state data and the addition of an internal feedback network increases the ability of the network itself to process dynamic information.
Setting $n$ as input and $m$ as output, the hidden layer and the following layer have $R$ neurons, the weight of the input layer to the hidden layer is $w_1$, the weight of the following layer to the hidden layer is $w_2$, and the weight of the hidden layer to the output layer is $w_3$; $x(k - 1)$ is the input of the neural network, using $u(k)$ to represent the output of the hidden layer, using the $u_c(k)$ to present the output of the undertaking layer, use the $y(k)$ to present the output of the neural network.

$$u(k) = f(w_2u_c(k) + w_1(x(k - 1)))$$ (7)

In formula

$$u_c(k) = u(k - 1)$$ (8)

$f$ is the transfer function of hidden layer and commonly used S type function

$$f(x) = (1 + e^{-x})^{-1}$$ (9)

g is the transfer function of output layer, and it is often a linear function

$$y(k) = g(w_3u(k))$$ (10)

Elman neural network uses the BP algorithm for weight correction, and the network error is

$$E = \sum_{k=1}^{m} (t_k - y_k)^2$$ (11)

t_k is the target output vector.

**Genetic algorithm**

Genetic algorithm is a highly parallel, random, and adaptive search algorithm developed by referring to natural selection and evolution mechanism. It uses a group search technique to represent a set of problems using a population. Through a series of genetic operations such as selection, crossover, mutation, and so on, it can produce a new generation of population and gradually make the population evolve into the state of the approximate optimal solution.

**Construction of GA-Elman algorithm**

When the genetic algorithm optimizes the connection weight of the neural network, some penalty items are added to the error function, and we have not to consider whether it is differentiable, which can improve the generality of the network and reduce the complexity of the network, so the optimization of the connection weights has great potential in this area. In optimizing the network structure, it can optimize different topologies in the face of diverse tasks.
to improve the insufficiency of the number of hidden neurons in the Elman network.

This study provides a method to optimize the connection weights and structure of neural networks. The main contents of the GA-Elman algorithm are chromosome encoding, definition of fitness function, and genetic operator construction, and the algorithm flow is shown in Figure 6.

Coding scheme. In the GA-Elman design process, the biggest problem encountered is the way to design a reasonable program to encode the network connection weights, output thresholds, and the number of hidden neurons. In this article, we adopt the methods of separate encoding and simultaneous evolution.

The number of hidden neurons is coded with real number and the upper limit $p$ is set. In addition, a binary code to the hidden layer is added as controlling gene which can be generated by a random function. When the value of the control gene is 0, corresponding to the hidden layer node has no effect on the output layer; when the value of the control gene is 1, it is equivalent to this hidden layer node, which influences the output layer. For the input $n$ and output $m$ of Elman neural network, its coding length is as follows

$$L = n \times p + p \times m + p + m = p \times (n + m + 1) + m$$

(13)

In formula, $n \times p$ indicates the length of the weight of the input layer to the hidden layer, $p \times m$ indicates the length of the weight of the hidden layer to the output layer, $p$ is the threshold coding length of hidden neurons, and $m$ is the threshold encoding length of neurons in the output layer. Consider the threshold in the weight code. Thus, this study is a piece of evolution together with a threshold.

Fitness calculation. We use the training error of the network to determine the fitness of the chromosome corresponding to the network and set fitness function as

$$F = C - e$$

(14)

$C$ is a constant and $e$ is the error. It is generally accepted that a network with a large error has a low degree of fitness.

Genetic operator construction. The selection operator uses “roulette” selection, namely, the probability of everyone entering the next generation is equal to the ratio of that adaptive value and total individual adaptation value of whole population. The higher the fitness value, the greater the probability of entering the next generation.

Crossover operator, using the “single point crossover”, partly exchanges or recombines the parental genes according to a certain probability.

Mutation operator binary coding uses the “inverted” mutations, and real coding is Gaussian.

Using the elite reservation strategy, individuals that searched currently have the highest degree of fitness, which will reserve to the next generation directly and prevent the loss or destruction of the crossover and mutation operations, thereby to ensure the global convergence of the genetic algorithm.

**GA-Elman recognition algorithm**

In this study, the optimization of both connection weights and hidden layer structures is performed simultaneously to maximize the optimization of Elman neural networks. According to the structure of the “Obscured fruit
recognition” section, a new optimization method is used to code the simultaneous evolution of the two methods. That is, while learning the structure of the hidden layer neurons, the connection weights can be learned. The new idea not only improves the network’s operating speed and recognition accuracy but also improves the generalization ability of the network to obtain a more accurate network model.

In this article, the basic steps of the optimization algorithm are as follows:

Step 1. Unifying the weights and thresholds of Elman network and, at the same time, coding the real number, the number of hidden layer neurons is coded with real number and the upper limit is set, adding the binary coded control gene to the hidden layer neurons;
Step 2. Decoding the one code, we can get different networks;
Step 3. The network from step 2 is trained by a given training to determine whether to meet the accuracy requirements, if satisfied, the training stops and turns to step 7, otherwise it transfers to step 4;
Step 4. According to the objective function and the training results to determine individual fitness, we use the elite retention strategy and select a number of individuals with the largest fitness value, directly inherited to the next generation;
Step 5. Using crossover and mutation operators to deal with the current generation of groups to generate the next generation of groups;
Step 6. Determining whether the set of genetic algebra is reached, then return to step 3, otherwise return to step 2;
Step 7. The optimization of Elman neural network model is obtained.

The feature of the extracted segmented image is modeled according to the above algorithm steps to realize the recognition of apple fruit, increase its recognition efficiency and precision, and further improve the harvesting efficiency of the apple harvesting robot.

Experiments

Experimental design and environment

In Fengxian, the China’s township of apple, 400 images are captured from apple orchard, and they are divided into training set and testing data according to 80/20, that is, 320 images as training samples and 80 images as training samples. In order to avoid the influence of the branches and leaves, the training samples also consider with the branches and leaves.

In order to better verify the superiority of the optimized algorithm, two groups experiments are arranged. One is compared with BP, Elman neural network, to show the performance of new algorithm optimized by GA. One is compared with some obscured fruit recognition algorithms, because single unobscured fruit recognition is no longer a challenge.

Test operation platform: CPU Intel Core2 Duo E7300 2.66 GHz, RAM 1.99 GB, Graphics card Intel® G33/G31 ECF; operating environment: 32 bit Windows 7, MatlabR2012b.

Recognition effect

Parameter setting. In training process, there are three indicators to measure the learning ability of each recognition model, such as training success rate, convergence, and run time. In the testing process, the cognitive performance of the trained model is evaluated from three aspects: recognition rate, recognition time, and error. According to different postures of target apples, the recognition rate of non-mask fruits, overlapped fruits, and fruits blocked by leaves and branches was calculated, respectively, and the total recognition rate was determined, to verify the generalization capabilities of three network models.

Genetic algorithm parameter setting: population size is 30, GA maximum iteration number is 300 times, the crossover rate is 0.9, the mutation rate is 0.01, the elite retention strategy selects four optimal individuals, the target error is MSE 1e-5, Elman is 2000 times, and traingdx is used as learning algorithm (a gradient descent learning algorithm with adaptive learning rate). Using Gao’s empirical formula, the number of neurons in hidden layer BP neural network is set

$$s = \sqrt{0.43mn + 0.12n^2 + 2.54m + 0.77n + 0.35 + 0.51}$$

(15)

where $s$ is the number of hidden neurons, $m$ is the input, and $n$ is the output.

Neural network training. In the 320 training samples, apple fruits, integrity leaves, and branches were extracted from the segmented images. There are 483 apples in 320 training images, and 50 leaves and 50 branches are selected to participate in training, simultaneously, thus 583 samples were considered as the training set. The extracted 16 features of each sample are used as the inputs of the neural network, and the outputs is three classes. So the neural network structure with sixteen input and three output is designed. Three kinds of neural networks for 100 times training, the statistical training success rate, three steps of convergences, training time, and the average values were listed in Table 1.

From Table 1 shows that the traditional Elman neural network is slightly higher than the BP network in training success rate; after optimizing the genetic algorithm, GA-Elman model’s training success rate greatly increased, and it can reach 100%. For the steps of convergence, the Elman
model is far lower than the BP network, indicating that the Elman model has better convergence performance. GA-Elman converged quickly and the optimization of genetic algorithm accelerated the convergence speed. From the training time, the shortest is Elman neural network. While GA-Elman algorithm takes the longest, because genetic optimization iteration time cost more, that is, by the time, cost is exchanged for training success ratio.

**Recognition and verify.** GA-Elman model was used as a classifier of the apple fruit to separate three types of fruit images from the background in terms of recognition effect. Then, a series of subsequent operations, such as binarization, small area denoising, contour extraction, and the boundary fitting, were carried out to obtain the final recognition effect, as shown in Figure 7. If the fruit was overlapped or obscured by the branches and leaves, we wouldn’t get the complete feature information. Then, the information will be lost, which will influence classification effect. In serious cases, misunderstandings or omissions may occur, as shown in Figure 8.

For operating efficiency and recognition rate of three models, the three models are evaluated from the following two aspects: one is recognition rate of fruits with different growth postures, the other is counting the total recognition rate, running time, and error. The recognition rate is

\[
\text{Recognition rate} = \frac{\sum \text{Correct samples}}{\text{Samples} \times \text{Experiment times}} \quad (16)
\]

Running time is the cost of each recognition operation, and its average value is recorded. The squared sum error between the output value of the model and the true value of

| Model      | Success rate (%) | Steps  | Time (s) |
|------------|------------------|--------|----------|
| BP         | 91               | 809 ± 97 | 2.297 ± 0.455 |
| Elman      | 94               | 519 ± 57  | 1.589 ± 0.326  |
| GA-Elman   | 100              | 300 ± 160 ± 33 | 5.031 ± 0.099  |

GA: genetic algorithm; BP: back-propagation.

**Figure 7.** Sample figure of correctly recognize fruit. (a) original image, (b) preliminary identification, (c) binarization, (d) small area denoising, (e) contour extraction, (f) fitting. I an unobscured fruit, II a slight shade, III a partial shade, IV an overlapping fruit.
the sample is calculated according to formula 11. The test was repeated 100 times and the results are shown in Table 2.

There are 139 apples in 80 testing images, 50 fruits were unobscured, 36 were overlapped, and 53 were obscured by branches or leaves, with recognition rate of images under different growth postures.

From Table 2, we can see that the recognition effect of the three identification models on unobscured fruits is more than 99%, so it is feasible to use the neural network algorithm in apple fruits recognition. Due to the presence of interference factors, we can’t recognize the lapped fruits well and the correct recognition rate slightly decreased. GA-Elman model of cognition’s effect is the best and its average recognition rate of lapped fruits and covering fruits can reach 88.67% and 93.64%, which were higher than the not-optimized traditional Elman identification model.

By evaluating the total recognition rate, model BP was found to be the lowest and model GA-Elman the highest, reaching 94.88%. From the sum of squared errors, the new model still performs optimally. GA-Elman NN model has the highest recognition and the least error, indicating that the model has the best identification accuracy. From the recognition time evaluation, the three recognition models are relatively close. Relatively speaking, the GA-Elman model is the shortest, which only needs 1.89 s to identity 179 samples of apple.

From the above experimental results, we came to know that three kinds of network algorithms can be trained for the training samples of apple image from the network model training process. From the recognition effect, for the single unobscured fruit, the recognition rates all of this three models are more than 99%, thus the neural network recognition algorithm is suitable for apple fruit recognition. The overall performance of the Elman neural network is better than BP neural network model either from the training performance of the network or from the recognition performance, because the Elman network has one more undertaking layer with respect to the BP network in structure and plays a memory role.

From the perspective of the training process of three kinds of algorithms, the training success rate of GA-Elman neural network is the highest, but the training takes long time. This demonstrates that genetic optimization is an iterative process at the expense of time, it can accelerate the convergence of the neural network, and save the time that artificially attempt network architecture. From the recognition effect, the recognition rate of the fruits with overlap and branches and leaves is improved greatly, the total recognition rate of the three types of fruits is also significantly increased, with smaller error. Therefore, the GA-Elman neural network model has the highest recognition accuracy.

### Obscured fruit recognition

In the current target fruit recognition algorithm of apple image, it is no longer a difficult problem to recognize a single unobscured target fruit. However, it is still very difficult to recognize obscured by apples, leaves, or branches, and the recognition rate is not high enough. In this experiment, two recognition algorithms of obscured apples are compared. The experimental data sets are 36 overlapped apples and 53 obscured apples, and the test was repeated 100 times and the results are presented in Table 3.

![Figure 8. Sample figure of missing recognize fruit. (a) original image, (b) preliminary identification, (c) binarization, (d) small area denoising, (e) contour extraction, and (f) fitting.](image)

| Table 2. Comparison of the recognition performance of each algorithm. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Model           | Unobscured fruit | Overlapped fruit | Obscured fruit  | Total recognition rate | Recognition time (s) |
| BP              | 99.54           | 85.61           | 88.85          | 92.36                | 2.20 ± 0.117     |
| Elman           | 99.68           | 86.86           | 91.42          | 93.02                | 1.97 ± 0.612     |
| GA-Elman        | 99.84           | 88.67           | 93.64          | 94.88                | 1.89 ± 0.183     |

GA: genetic algorithm; BP: back-propagation.

Jia et al. 11
to a large extent. The statistics of 118 overlapping and obscured fruits were carried out. When the area covered by the fruit was less than 30%, the recognition rate has almost no effect; and the covered area was between 30% and 60%, the recognition effect was affected, and it was easy to cause errors. When the area covered is greater than 60%, or even larger, identifying the difficulties caused by the appearance of a greater chance of misunderstanding or leakage identification could not be recognized.

Results analysis

According to the results of the first group experimental, it can be known from the network model training process that three network algorithms can be trained for training samples of apple images. From the recognition effect point of view, for the single unobscured fruit, the recognition rate of the three models is greater than 99%, so the neural network recognition algorithm is suitable for apple fruit recognition. Comparing the training processes of these three algorithms, GA-Elman neural network has the highest training success rate, but the training time is longer. This shows that genetic optimization is an iterative process at the cost of time, which can accelerate the convergence of neural networks and save the time of manually attempting the network structure. Judging from the recognition effect, the recognition rate of fruits obscured by fruits or branches and leaves has been greatly improved, and the total recognition rate of the three types of fruits has also been significantly improved with less error. Therefore, GA-Elman neural network model has the highest recognition precision.

Recognition of obscured and overlapped apples is still a difficult problem; from second group experimental results, our method is better than the other two methods. However, when the area is too large to be covered by branches and leaves, the extraction of apple fruit is not complete, and the loss of information is relatively large, which leads the sample to be the outlier samples. Outlier samples are still difficult to machine learn, and so how to optimize the outlier samples will be the focus of the next research, thereby improving the recognition efficiency.

Conclusions and further works

PCNN is applied in apple image segmentation. Apple fruit, leaves, and branches got that based on the image space approach and the gray level similarity, the segmentation effect is good and reaches the expectation. As for split image, six color features were extracted from R, G, B, H, S and I channels respectively, Hu invariant moment these 10-feature vectors. The extracted 16 feature vectors can describe the global and local features of the object well, and the extracted features are effective. After the Elman neural network is optimized by the genetic algorithm, it accelerates the network convergence, improves the generalization ability of the network, and greatly improves the training success rate, which are reach 100% in this study. Only genetic optimization pays the price of time, and network learning time is slightly longer. Neural network models have high recognition efficiency for unobscured apples, which is close to 100%, and the neural network for recognition of apple fruit is feasible. For the recognition of overlapped apples, the correct recognition rate of the GA-Elman optimization model is significantly higher than the other models. It shows that the new GA-Elman model improves the efficiency of network operation and the highest recognition accuracy.

In conclusion, GA-Elman algorithm will consume more time; however, in the training process, the convergence speed will be improved, not only training success rate is increasing, but also generalization ability is enhancing. It can reduce the time of artificial attempts. The recent literatures about real-life applications of contemporary automation techniques in different fields give us new inspiration, using other intelligent computation39–42 to optimize neural network, in order to get more superior model; at the same time, we try to use deep learning algorithm43–46 to recognize the apple target fruit. We expect that the new algorithm can be applied in more fields, such as target recognition,47–50 sliding control,51–54 and so on.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by Natural Science Foundation of Shandong Province in China (no.: ZR2017BC013); China Postdoctoral Science Foundation (no.: 2018M630797); National Nature Science Foundation of China (no.: 31571571, 61572300, 61379101); and Taishan Scholar Program of Shandong Province of China (no.: TSHW201502038).

ORCID iD

Weikuan Jia  https://orcid.org/0000-0001-6242-3269

References

1. Zhao DA, Lv JD, Ji W, et al. Design and control of an apple harvesting robot. Biosyst Eng 2011; 110(2): 112–122.
2. Baeten J, Donné K, Boedrij S, et al. Autonomous fruit picking machine: a robotic apple harvester. *Field Robot Ser Robot* 2008; 42: 531–539.

3. Ferrara P, Piva A, Argenti F, et al. Wide-angle and long-range real time pose estimation: a comparison between monocular and stereo vision systems. *J Vis Commun Image R* 2017; 48: 159–168.

4. Wei J, Zhijie Q, Bo X, et al. A nighttime image enhancement method based on Retinex and guided filter for object recognition of apple harvesting robot. *Int J Adv Robot Syst* 2018; 15(1). DOI: 10.1177/1729881417753871.

5. Gongal A, Karkee M, and Amatya S. Apple fruit size estimation using a 3D machine vision system. *Inf Process Agri* 2018; 5(4): 498–503.

6. Lv JD, Zhao DA, and Ji W. Research on matching recognition method of oscillating fruit for harvesting robot (in Chinese). *Trans Chin Soc Agri Eng* 2013; 29(20): 32–39.

7. Liu X, Zhao D, Jia W, et al. A detection method for apple fruits based on color and shape features. *IEEE Access* 2019; 7: 67923–67933.

8. Song HB, Zhang CD, Pan JP, et al. Segmentation and reconstruction of overlapped apple images based on convex hull (in Chinese). *Trans Chin Soc Agri Eng* 2013; 29(3): 163–168.

9. Zhang CL, Zhang J, Zhang JX, et al. Recognition of green apple in similar background (in Chinese). *Trans Chin Soc Agri Mach* 2014; 45(10): 277–281.

10. Ruan CZ, Zhao DA, Jia WK, et al. Night vision image denoising of apple harvesting robots based on the wavelet fuzzy threshold. *Int J Adv Robot Syst* 2015; 12(12): 169.

11. Kelman EE and Linker R. Vision-based localisation of mature apples in tree images using convexity. *Biosyst Eng* 2014; 118: 174–185.

12. Wachs JP, Stern HI, Burks T, et al. Low and high-level visual feature-based apple detection from multi-modal images. *Precis Agri* 2010; 11(6): 717–735.

13. Li PL, Lee SH, and Hsu HY. Review on fruit harvesting method for potential use of automatic fruit harvesting systems. *Procedia Eng* 2011; 23: 351–366.

14. Sui XD, Zheng YJ, Wei BZ, et al. Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks. *Neurocomputing* 2017; 237: 332–341.

15. Deng X, Zheng YJ, Xu YL, et al. Graph cut based automatic aorta segmentation with an adaptive smoothness constraint in 3D abdominal CT images. *Neurocomputing* 2018; 310: 46–58.

16. Bulanon DM, Kataoka T, Ota Y, et al. Segmentation algorithm for the automatic recognition of Fuji apples at harvest. *Biosyst Eng* 2002; 83(4): 405–412.

17. Huang LW, He DJ, and Yang SX. Segmentation on ripe fuji apple with fuzzy 2d entropy based on 2d histogram and GA optimization. *Intell Autom Soft Co* 2013; 19(3): 239–251.

18. Wang DD, Xu Y, Song HB, et al. Fusion of K-means and Neut algorithm to realize segmentation and reconstruction of two overlapped apples without blocking by branches and leaves (in Chinese). *Trans Chin Soc Agri Eng* 2015; 31(10): 227–234.

19. Mizushima A and Lu RF. An image segmentation method for apple sorting and grading using support vector machine and Otsu’s method. *Comput Electron Agri* 2013; 94: 29–37.

20. Subashini MM and Sahoo SK. Pulse coupled neural networks and its applications. *Exp Syst Appl* 2014; 41(8): 3965–3974.

21. Zhu Z and Liu P. Feasibility research of text information filtering based on genetic algorithm. *Sci Res Essays* 2010; 5(22): 3405–3410.

22. Kumar M, Husian M, Upreti N, et al. Genetic algorithm: review and application. *Int J Inf Technol Knowl Manag* 2010; 2(2): 451–454.

23. Li H, Zhang G, and Zhu Z. Color image segmentation based on PCNN. *J Math Inf* 2018; 13: 41–53.

24. Wang H, Wu J, Zhang P, et al. Learning shapelet patterns from network-based time series data. *IEEE T Ind Inf* 2019; 15(7): 3864–3876.

25. Ji C, Zhao C, Liu S, et al. A fast shapelet selection algorithm for time series classification. *Comput Netw* 2019; 148: 231–240.

26. Thyagarajan KK and Kalaiarasi G. Pulse coupled neural network based near-duplicate detection of images (PCNN-NDD). *Adv Electr Comput Eng* 2018; 18(3): 87–97.

27. Shi D, Zhu L, Cheng ZY, et al. Unsupervised multi-view feature extraction with dynamic graph learning. *J Vis Commun Image R* 2018; 56: 256–264.

28. Mei QL, Zhang HK, and Liang C. A discriminative feature extraction approach for tumor classification using gene expression data. *Curr Bioinform* 2016; 11(5): 561–570.

29. Li J, Xu WT, Wan WB, et al. Movie recommendation based on bridging movie feature and user interest. *J Comput Sci* 2018; 26: 128–134.

30. Ji W, Meng XL, Qian ZJ, et al. Branch localization method based on the skeleton feature extraction and stereo matching for apple harvesting robot. *Int J Adv Robot Syst* 2017; 14(3): 172.

31. Wang H, Zhang P, Zhu X, et al. Incremental subgraph feature selection for graph classification. *IEEE T Knowl Data Eng* 2016; 29(1): 128–142.

32. Luo C, Tan C, Wang X, et al. An evolving recurrent interval type-2 intuitionistic fuzzy neural network for online learning and time series prediction. *Appl Soft Comput* 2019; 78: 150–163.

33. Zhang H, Ji H, and Wang X. Transfer learning from unlabeled data via neural networks. *Neural Process lett* 2012; 36(2): 173–187.

34. Bhusan B, Singh M, and Hage Y. Identification and control using MLP, Elman, NARXSP and radial basis function networks: a comparative analysis. *Artif Intell Rev* 2012; 37(2): 133–156.

35. Jia WK, Zhao DA, Zheng YJ, et al. A novel optimized GA-Elman neural network algorithm. *Neural Comput Appl* 2019; 31(2): 449–459.
36. Kaikhah K and Garlick R. Variable hidden layer sizing in Elman recurrent neuro-evolution. *Appl Intell* 2000; 12(3): 193–205.

37. Gao DQ. On structures of supervised linear basis function feed forward three-layered neural networks (in Chinese). *Chin J Comput* 1998; 21(1): 80–86.

38. Jia WK, Zhao DA, Hu CL, et al. Fast recognition of overlapping fruit based on maximum optimization for apple harvesting robot. *Int J Collab Intell* 2015; 1(2): 124–136.

39. Liu H, Xu B, Lu DJ, et al. A path planning approach for crowd evacuation in buildings based on improved artificial bee colony algorithm. *Appl Soft Comput* 2018; 68: 360–376.

40. Yang Q, Chen WN, Yu Z, et al. Adaptive multimodal continuous ant colony optimization. *IEEE T Evolut Comput* 2016; 21(2): 191–205.

41. Zhai LB and Wang H. Crowdsensing task assignment based on particle swarm optimization in cognitive radio networks. *Wirel Commun Mob Comput* 2017; 2017: 1–9.

42. Lv C, Jiang W, Hu S, et al. Efficient dynamic evolution of service composition. *IEEE T Ser Comput* 2015; 11(4): 630–643.

43. LeCun Y, Bengio Y, and Hinton G. Deep learning. *Nature* 2015; 521(7553): 436.

44. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015; 61: 85–117.

45. Lian J, Hou SJ, Sui XD, et al. Deblurring retinal optical coherence tomography via a convolutional neural network with anisotropic and double convolution layer. *IET Comput Vis* 2018; 12(6): 900–907.

46. Zhang B, Zhu L, Sun JD, et al. Cross-media retrieval with collective deep semantic learning [J]. *Multimed Tools Appl* 2018; 77: 22247–22266.

47. Hou S, Zhou S, Liu W, et al. Classifying advertising video by topicalizing high-level semantic concepts. *Multimed Tools Appl* 2018; 77(19): 25475–25511.

48. Sun J, Wang Y, Li J, et al. View-invariant gait recognition based on kinect skeleton feature. *Multimed Tools Appl* 2018; 77: 24909–24935.

49. Witus IK, On CK, Alfred R, et al. A review of computer vision methods for fruit recognition. *Adv Sci Lett* 2018; 24(2): 1538–1542.

50. Liu B, Liu H, Zhang H, et al. A social force evacuation model driven by video data. *Simul Model Pract Th* 2018; 84: 190–203.

51. Liu H, Liu BX, Zhang H, et al. Crowd evacuation simulation approach based on navigation knowledge and two-layer control mechanism. *Inform Sci* 2018; 436–437: 247–267.

52. Ding S, Chen WH, Mei K, et al. Disturbance observer design for nonlinear systems represented by input-output models. *IEEE Trans Ind Electron* 2019; 67(2): 1222–1232. DOI: 10.1109/TIE.2019.2898585.

53. Ríos H, Falcón R, González OA, et al. Continuous sliding-mode control strategies for quadrotor robust tracking: real-time application. *IEEE Trans Ind Electron* 2018; 66(2): 1264–1272.

54. Zhang P, Liu H, and Ding Y H. Crowd simulation based on constrained and controlled group formation. *Vis Comput* 2015; 31(1): 5–18.