



# Farmland Weed Species Identification Based on Computer Vision

Shengping Liu<sup>1</sup>, Junchan Wang<sup>2</sup>, Liu Tao<sup>3</sup>, Zhemin Li<sup>1</sup>,  
Chengming Sun<sup>3</sup>, and Xiaochun Zhong<sup>1</sup>✉

<sup>1</sup> Institute of Agriculture Information,  
Chinese Academy of Agriculture Sciences/Key Laboratory of Agro-Information  
Services Technology, Ministry of Agriculture, Beijing, China  
{liushengping, lizhemin, zhongxiaochun}@caas.cn

<sup>2</sup> Lixiahe Regional Institute of Agricultural Sciences of Jiangsu Province,  
Yangzhou, China  
wjc@wheat.org.cn

<sup>3</sup> Jiangsu Key Laboratory of Crop Genetics and Physiology/Co-Innovation  
Center for Modern Production Technology of Grain Crops,  
Yangzhou University, Yangzhou, China  
{tliu, cmsun}@yzu.edu.cn

**Abstract.** In order to alleviate the difficulties in collecting indexes for the analysis of farmland weed communities, we implemented a computer vision technology-based method for the identification of farmland weeds at the species level. By using the super-green and maximum interclass difference methods to obtain a green vegetation binary image, we were able to separate weeds from cultivated crops through multiple etching and the removal of small areas. A BP (back propagation) neural network was used for weed recognition, and the morphological characteristics of the weeds and each region were selected following etching to construct the input matrix of the recognition model for training and testing the BP network. After experimenting with the computational vision method for the identification of five weed species, we discovered that the recognition accuracy rate reached 96%. The results showed that the computer vision method could quickly and accurately extract a weed community analysis index, thereby providing a reference for the intelligent analysis of weed communities.

**Keywords:** Weed communities · Index extraction · Image processing  
Class · Computer vision

## 1 Introduction

Rice paddy weeds compete with rice for sunlight, nutrients, and water, which can severely impact the yield and quality of rice [1, 2]. Conversely, weeds are an important component of a balanced and stable farmland ecosystem and play an important role in holding soil moisture and fertilizers, accelerating nutrient cycling, and providing a living environment for soil organisms [3, 4]. Studying farmland weed communities is therefore of considerable significance for ensuring food security and maintaining ecosystem stability.

Weed species constitute the primary quantitative index in studying farmland weed communities. Three primary methods for identifying weeds in the field exist: the manual identification method, the remote sensing method, and the computer vision-based recognition method. Currently, data in a fixed area is mainly obtained manually [5–8]. The investigation of weed species usually requires the professional evaluation or the assistance of the weed primary color spectrum [3].

The development of computer vision and image processing technologies has played an important role in weed species identification [12–17]. The identification of weed species has mainly been performed by extracting morphological, color, and texture parameters, and using these characteristic parameters to establish a weed identification model. The identification models commonly used include the camera supported model and the artificial neural network model [16–19].

Some scholars have used leaf color to identify weeds at the species level, but the recognition speed and accuracy is low and requires further improvement. Additionally, some scholars use morphological characteristics in identification, however, extracting the morphological characteristics is difficult as the leaves tend to intertwine with each other in the natural environment, and thus this method is not conducive for outdoor weed identification. Therefore, accurate weed identification cannot be achieved using only one factor, and multiple features are required for a comprehensive analysis. In this study, by extracting characteristic morphological, color, and texture parameters, an identification model for weed species was established using multi-characteristic parameters. The current research mainly uses a computer vision system to identify weeds from crop plants and measure the soil background to calculate their distribution [16, 19]. Few studies have focused on the identification of farmland weed communities. In this study, we developed a computer vision extraction method for weed species, thereby providing a reference for the intelligent analysis of farmland weed communities.

## 2 Materials and Methods

The images used in this study were obtained from rice-wheat rotation fields from 2014 to 2016. Five common species of field weeds were selected for the species recognition experiments. There were 40 images in total ( $750 \times 750$  pixels) for each weed species. The identification process is shown in Fig. 1.

### 2.1 Extraction

As shown in Fig. 2a, the original image includes rice plants, *Monochoria* seedlings, duckweed, and *Spirogyra*. The green vegetation of the samples was extracted using Eq. (1) and the OTSU method [17], and the extraction results are shown in Fig. 2b (including crops, weeds, and *Spirogyra*). In the RGB images (Fig. 3), the green band of *Spirogyra* was relatively smaller than the weeds and crops. Further threshold segmentation using green bands reduced the effects of the *Spirogyra* on weed extraction. The results are shown in Fig. 2c. The method was effective at reducing the influence of *Spirogyra* in paddy fields.

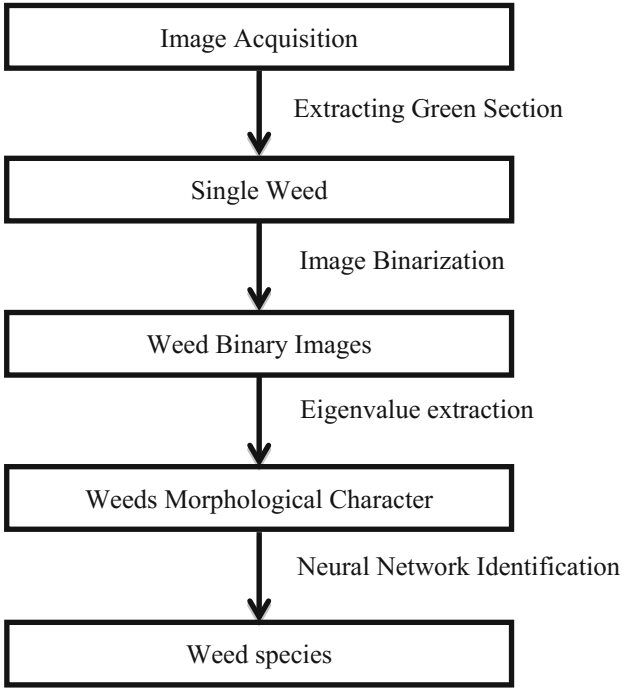


Fig. 1. The processes of identify weed types

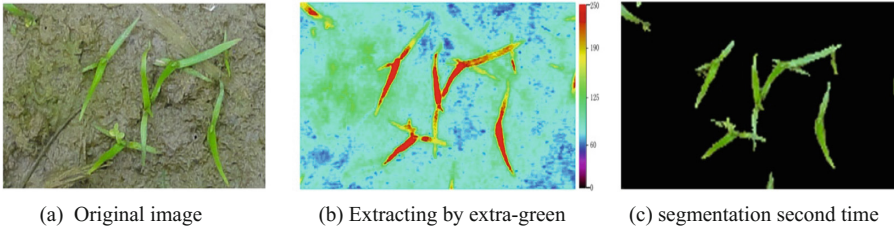
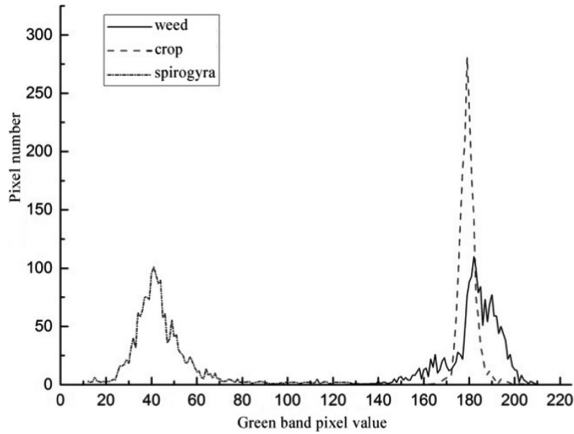


Fig. 2. Extracting vegetation (Color figure online)

$$ExG = 2G - R - B \tag{1}$$

We first removed the areas with fewer than 1,000 pixels in Fig. 4a, and then used a  $3 \times 3$  template to etch the binary images, and then removed the areas with fewer than 800 pixels in the binary images. Following this, we used the same operation to etch the

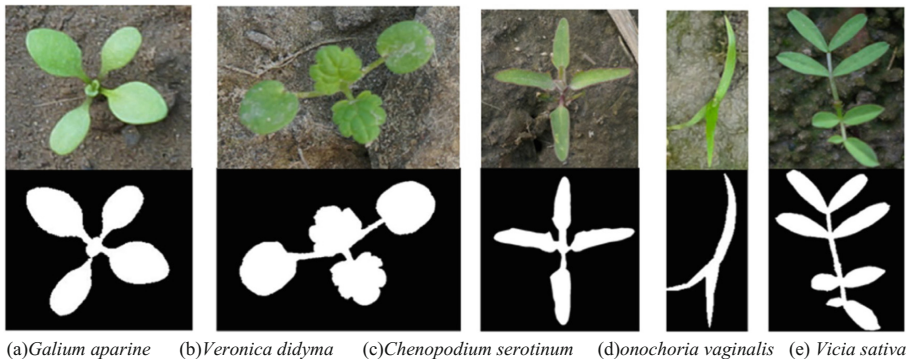


**Fig. 3.** Green band histogram of crop/weed/spirogyra

binary images twice and removed the areas with fewer than 600 pixels and 400 pixels; performed 3 expansion operations on the obtained binary images by using the  $3 \times 3$  templates; filled the smaller holes in the region by using the hole-filling algorithm [21]; and obtained the binary images of the weeds.

## 2.2 Weed Species Identification

Weeds are difficult to identify at the seedling stage. We therefore selected *Galium aparine*, *Veronica didyma*, *Chenopodium serotinum*, *Monochoria vaginalis*, and *Vicia sativa* as study objects. The binary images of the weeds (Fig. 4) were obtained by using the green vegetation extraction method detailed in Sect. 2.1.



**Fig. 4.** Binarization for weed images

### 2.3 Feature Parameter Selection

Due to the influence of light, image acquisition equipment, and other factors, the color and texture characteristics of the same weed species obtained in the wild often differ. However, the morphological characteristics are not affected by these environmental factors. Therefore, we selected the following morphological parameters to construct the feature vector for identifying weed species:

compactness:

$$\text{compactness: } Co = \frac{4\pi A}{P} \quad (2)$$

$$\text{quasi circularity: } R = \frac{4A}{\pi M} \quad (3)$$

$$\text{foliaceous: } B = \frac{R}{M} \quad (4)$$

$$\text{centrifugal rate: } B = \frac{S}{L} \quad (5)$$

Where A is the area of the binary region, P is the circumference of the binary region, M is the major axis length of the smallest circumscribed rectangle, R is the shortest distance from the regional centroid to the boundary, S is the minor axis length of the region, and L is the major axis length of the region. To better reflect the blade shape and blade configuration, we used a template of a certain size (the size of the template approximately equal to 1/60 of the binary region) to etch the region. The number *N* of the binary region in the image following etching is the characteristic value, and the mean and variance of the quasi circularity, foliaceous, and centrifugal rate of each region also constitute characteristic values used to identify the weed species.

### 2.4 Selection of the Recognition Model

In this study, we used the BP (back propagation) neural network to identify the five weed species. The back propagation (BP) neural network is a multi-layer feedforward network trained according to error back-propagation algorithm and is one of the most widely used neural network models. BP network can be used to learn and store a large number of mapping relations of input-output model without prior description of the mapping relationship [22]. In the study, 20 images were selected for each species for the construction and training of the network. The network input data constitute the 11 characteristic values described in Sect. 1.3 and were extracted from each image. The input data were normalized to increase the network convergence speed and diagnostic accuracy [23]. A three-layer BP neural network was constructed, and the network input

layer consisted of 11 neurons corresponding to 11 eigenvalues, respectively. There were 20 neurons in the hidden layer, and one neuron in the output layer. The learning rate was set to 0.5, the inertial coefficient was 0.8, the number of iterations was set to 1,000, and the target error was set to 0.001. The hidden layer transfer function was selected as the Tansig type, namely:

$$f(x) = \frac{2}{1 + e^{-ax}} - 1 \tag{6}$$

Where a is the slope function of the Tansig function, and a is 2 in this paper.

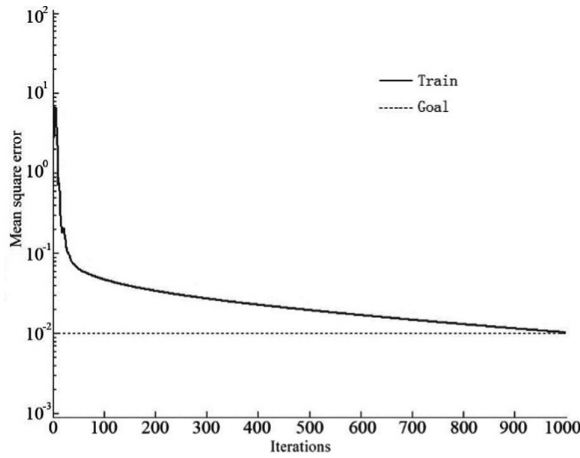


Fig. 5. Training iteration of weed identification

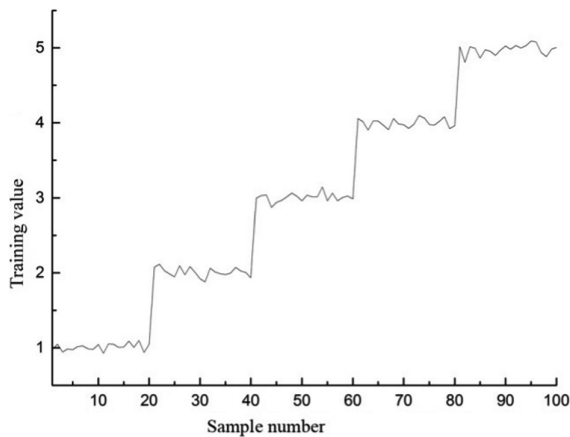


Fig. 6. Simulation output of training samples

The training process of the weed recognition model is shown in Fig. 5. The 100 network training samples were numbered in sequence: 1–20 for *G. aparine*, 21–40 for *V. didyma*, 41–60 for *C. serotinum*, 61–80 for *M. vaginalis*, and 81–100 for *V. sativa*. The input sample matrix of the network training was  $11 \times 100$  and the output matrix of the network training was  $1 \times 100$ , with each element corresponding to the classification of the sample and a value of 1, 2, 3, 4, and 5, respectively. After 1,000 iterations, the target error of 0.001 was achieved. The training results are shown in Fig. 6. It was evident that all the samples had been correctly identified.

### 3 Results

The 100 samples for the verification of the recognition model were numbered in sequence: 1–20 for *G. aparine*, 21–40 for *V. didyma*, 41–60 for *C. serotinum*, 61–80 for *M. vaginalis*, and 81–100 for *V. sativa*. We extracted the 11 characteristic values of each sample to construct the  $11 \times 100$  input matrix and normalized the data. The normalized test samples were input into the network to perform the simulation. The

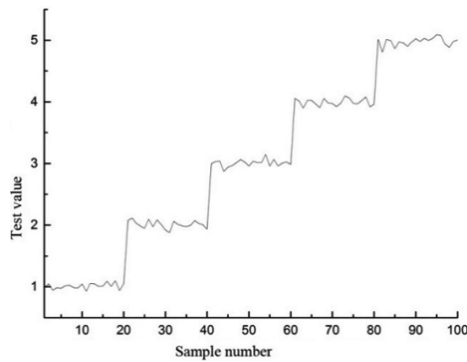


Fig. 7. Simulation output of test examples

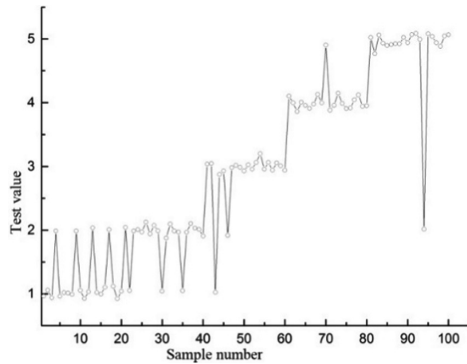


Fig. 8. Simulation output of test examples (only use characteristics of weeds itself)

output result of the simulation is shown in Fig. 7, and all test samples were identified. The recognition accuracy was only 87% when only the four feature values for the weeds were used in the training and development of the method, without using the morphological characteristics of each region following etching. The recognition was even lower for *G. aparine* and *V. didyma* at only 80% and 85%, respectively. The simulation output result is shown in Fig. 8.

## 4 Discussion

Common methods of farmland weed recognition include these features: threshold segmentation method based on machine vision, weed identification method based on genetic algorithm, weed identification method based on polarization spectrum and weed identification method based on BP network. According to the external characteristics of weeds, threshold segmentation method based on machine vision realizes image segmentation and weed recognition through using the threshold segmentation model. The recognition rate was about 85% [24]. The weed recognition method based on genetic algorithm can identify weeds quickly and accurately, and reduce the occurrence of local optimal value. The recognition rate was about 93% [25]. The method of recognizing weeds based on polarization spectrum can identify the weed model well by analyzing and comparing the spectral response law, spectrum characteristics and decision precision of weeds, and the comprehensive recognition rate of different polarization states reaches more than 90% [26]. After the training of samples, weed recognition method based on BP neural network can recognize 96% of weeds in seedling stage. According to different site requirements, the most suitable identification algorithm can be selected for application in practical application.

## 5 Conclusion

The correct identification of weed species in farmlands is a prerequisite for the analysis of farmland weed communities. In this study, we proposed a weed extraction method for the field environment and successfully identified weed species in this environment using a combination of 11 morphological features and a BP neural network. We reached the following conclusions:

- (1) We proposed a computer vision-based weed extraction method for the field environment. The green vegetation in the field could be successfully extracted by using the super-green method and OTSU algorithm. The effect of *Spirogyra* on the paddy field images was eliminated by using the green component and the OTSU algorithm in the RGB mode.
- (2) We proposed a weed identification method based on the BP neural network. Including the morphological characteristics of the weeds and of each area following etching could better reflect the weed species. When only weed characteristic values were used, the recognition accuracy rate was 87%, and the accuracy rate did not exceed 85% for the identification of *G. aparine* and *V. didyma*, which



share similar morphological characteristics. The network trained by the two types of characteristic values at the same time was thus able to identify the samples. This indicated that the method proposed in this paper could effectively identify and classify field weeds with a recognition accuracy rate above 96%.

- (3) In this paper, we segmented and identified five common weed species occurring in the field environment. The process was neither affected by the external lighting conditions nor the image acquisition equipment, has wide adaptability, and can provide a reference for the recognition of other weeds under natural conditions. However, weeds at the seedling stage were selected for the weed identification experiment, and thus the identification of weeds at maturity still requires further study.

**Acknowledgment.** This research was supported by Technology Innovation Project Fund of Chinese Academy of Agricultural Sciences (2017).

## References

1. Shuren, C., Huadong, Z., Ruimei, W., et al.: Identification for weedy rice at seeding stage based on hyper-spectral imaging technique. *Trans. Chin. Soc. Agric. Mach.* **44**(05), 253–257 (2013)
2. Ali, A., Streibig, J.C., Andreasen, C.: Yield loss prediction models based on early estimation of weed pressure. *Crop. Prot.* **53**, 125–131 (2013)
3. Shi, L., Shen, M., Jiang, M.: Effect of long-term different fertilization management on weed community in rice-wheat rotation field. *Sci. Agric. Sin.* **46**(02), 310–316 (2013)
4. Lake, E.C., Hough-Goldstein, J., Shropshire, K.J., et al.: Establishment and dispersal of the biological control weevil *Rhinoncomimus latipes* on mile-a-minute weed, *Persicaria perfoliata*. *Biol. Control* **58**(3), 294–301 (2011)
5. Radicetti, E., Mancinelli, R., Campiglia, E.: Impact of managing cover crop residues on the floristic composition and species diversity of the weed community of pepper crop (*Capsicum annum* L.). *Crop Protection*. **44**, 109–119 (2013)
6. Cheng, C., Wan, K., Tao, Y., et al.: The effects of fertilization on weed communities and wheat growth in winter wheat (*Triticum aestivum* L.) field under different cropping rotations. *Ecol. Environ. Sci.* **22**(03), 370–378 (2013)
7. Dong, C., Liu, Q., Gao, J., et al.: Effects of different fertilization models on the characteristics of weed communities during the rice growing seasons. *Acta Prataculturae Sin.* **22**(03), 218–226 (2013)
8. Yang, R., Yongzhong, S.: Effects of cultivation regimes on weed community structures in newly reclaimed sandy farmlands. *Chin. J. Eco-Agric.* **18**(06), 1218–1222 (2010)
9. Dongjian, H., Yongliang, Q., Pan, L., et al.: Weed recognition based on SVM-DS multi-feature fusion. *Trans. Chin. Soc. Agric. Mach.* **44**(02), 182–187 (2013)
10. Zhang, W., Bingjun, L., Shi, W.: Determination of vegetation coverage by photograph and automatic calculation. *Bull. Soil Water Conserv.* **29**(02), 39–42 (2009)
11. Hu, L., Luo, X., Zeng, S., et al.: Plant recognition and localization for intra-row mechanical weeding device based on machine vision. *Trans. Chin. Soc. Agric. Eng.* **29**(10), 12–18 (2013)
12. Bauer, T., Strauss, P.: A rule-based image analysis approach for calculating residues and vegetation cover under field conditions. *CATENA* **113**, 363–369 (2014)

13. Swain, K.C., Nørremark, M., Jørgensen, R.N., et al.: Weed identification using an automated active shape matching (AASM) technique. *Biosyst. Eng.* **110**(4), 450–457 (2011)
14. Tellaeche, A., Pajares, G., Burgos-Artizzu, X.P., et al.: A computer vision approach for weeds identification through Support Vector Machines. *Appl. Soft Comput.* **11**(1), 908–915 (2011)
15. Zhang, J., Yang, H.: Application of self-organizing neural networks to classification of plant communitiee in Panquangou nature reserve, North China. *Acta Ecol. Sin.* **27**(03), 1005–1010 (2007)
16. Deyao, F., Qing, Y., Baojun, Y., et al.: Progress in research on intelligentization of field weed recognition and weed control technology. *Sci. Agric. Sin.* **43**(09), 1823–1833 (2010)
17. Zhou, J., Wang, M., Shao, Q.: Adaptive segmentation of field image for green plants. *Trans. Chin. Soc. Agric. Eng.* **29**(18), 163–170 (2013)
18. Zhang, D., Zhou, C., Zhou, Q., et al.: Hole-filling algorithm based on contour. *J. Jilin Univ. (Sci. Ed.)* **49**(1), 82–86 (2011)
19. Zou, X., Ding, W., Liu, D., et al.: Classification of rice planthopper based on invariant moments and BP neural network. *Trans. Chin. Soc. Agric. Eng.* **29**(18), 171–178 (2013)
20. Liu, X.: Study on data normalization in BP neural network. *Mech. Eng. Autom.* **03**, 122–123 (2010)
21. Pulido, C., Solaque, L., Velasco, N.: Weed recognition by SVM texture feature classification in outdoor vegetable crop images. *Ing. E Investig.* **37**(1), 68–74 (2017)
22. Pantazi, X.-E., Moshou, D., Bravo, C.: Active learning system for weed species recognition based on hyperspectral sensing. *Biosyst. Eng.* **146**(SI), 193–202 (2016)
23. Rahman, M., Blackwell, B.: Smartphone-based hierarchical crowdsourcing for weed identification. *Comput. Electron. Agric.* **113**, 14–23 (2015)
24. Saha, D., Hanson, A. and Shin, S.Y.: Development of enhanced weed detection system with adaptive thresholding and support vector machine. In: *Conference on Research in Adaptive and Convergent Systems (RACS)*. ACM, pp. 85–88 (2016)
25. Jianjiao, Y., Ji, Q., Fengwu, Z.: Study on identification of the field weed based on genetic neural network. *J. Chin. Agric. Mech.* **37**(9), 223–226, 230 (2016)
26. Fenfang, L., Dongyan, Z., Wang Xiu, W., Taixia, C.X.: Identification of corn and weeds on the leaf scale using polarization spectroscopy. *Infrared Laser Eng.* **45**(12), 361–370 (2016)