IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 2, 2020, accepted June 9, 2020, date of publication June 12, 2020, date of current version June 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3001999

Recognition of Weeds in Wheat Fields Based on the Fusion of RGB Images and Depth Images

KE XU[1,2,3,4], HUAIMIN LI[1,2,3,4], WEIXING CAO[1,2,3,4], YAN ZHU[1,2,3,4], RONGJIA CHEN[1], AND JUN NI[1,2,3,4]

[1]College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China
[2]National Information Agricultural Engineering Technology Center, Nanjing Agricultural University, Nanjing 210095, China
[3]Engineering Research Center of Smart Agriculture, Ministry of Education, Nanjing Agricultural University, Nanjing 210095, China
[4]Jiangsu Collaborative Innovation Center for the Technology and Application of Internet of Things, Nanjing Agricultural University, Nanjing 210095, China

Corresponding author: Jun Ni (nijun@njau.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFD0300606, in part by the National Natural Science Foundation of China under Grant 31871524, in part by the Six Talent Peaks Project in Jiangsu Province under Grant XYDXX-049, in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and in part by the Primary Research and Development Plan of Jiangsu Province of China under Grant BE2017385.

ABSTRACT Due to the low recognition rate of weeds in wheat fields and the inability to accurately locate weeds, we propose a recognition method for weeds in natural wheat fields based on the fusion of RGB image features and depth features. The method breaks through the limitations of the two-dimensional spatial features extracted from RGB images when recognizing grass weeds similar to wheat. According to the species, distribution of weeds in wheat fields, we extracted the color, position, texture, and depth features of weeds in wheat fields from RGB and depth images during the tillering and jointing stages. And then used the AdaBoost algorithm for the integrated learning of multiple classifiers, thereby achieving the recognition of weeds in wheat fields. The experimental results revealed that the recognition speed of weeds during the tillering stage was 0.2 s and the accuracy rate was 88%. The recognition speed of weeds during the jointing stage was 0.69 s, and the accuracy rate of weed recognition was 81.08%. These results are significantly higher than the weed recognition rate based on features extracted from RGB images.

INDEX TERMS Recognition of weeds in wheat fields, RGB-D images, depth features, AdaBoost.

I. INTRODUCTION

Weeds compete with wheat for light, water, fertilizer, and growth space [1]–[3], and have become major biological disasters that restrict wheat yield and quality. Among many weed control methods, chemical weeding has become the main approach used in field weeding due to its high efficiency [4], [5]. Due to the lack of weed species and distribution information, the use of chemical herbicides mostly consists of extensive spraying. Excessive application can cause serious environmental pollution [6], affect the yield and quality of agricultural products, and reduce agricultural production efficiency. In order to address the issue of extensive spraying, the study of site-specific weed management (SSWM) is very important [7], [8]. The recognition of weeds in wheat fields is a key step in the implementation of SSWM, and is also the core technology being used in the development of intelligent weeding robots.

The types of weeds in wheat fields can mainly be divided into grass weeds and broad-leaved weeds [8]. Grass weeds (Poaceae family) are monocotyledonous plants with long and narrow leaves, extremely similar to wheat. Broad-leaved weeds with relatively broad leaves are more different from the shape of wheat than grass weeds. In recent years, changes in farming methods and the frequent introduction of seeds in various wheat areas have led to great changes in the species and emergence degree of weeds in wheat fields. Grass weeds have invaded wheat fields and become the dominant populations, and, like broad-leaved weeds, their presence has become a prominent issue affecting wheat production [9].

Spectrum, image, and spectral imaging technologies are currently the main research methods for the rapid recognition of field weeds. The classification method of weeds based on spectral technology consists of recognizing differences in the reflectance of wheat, weeds, and soil within a certain

The associate editor coordinating the review of this manuscript and approving it for publication was Ramakrishnan Srinivasan.

band [10], [11]. Due to the similar leaf layer structure of grass weeds and wheat, however, the spectral features of reflectance at specific wavelengths are very much alike [12], making it difficult to distinguish the two via differences in spectral reflectance. The classification technology of weeds in wheat fields based on image processing consists of recognizing weeds in wheat fields by analyzing the color, position, shape, and texture of wheat and weeds in wheat field images [13], [14]. It is difficult, however, to distinguish grass weeds from the features extracted from two-dimensional images. Spectral imaging technology uses the comprehensive features of images to segment soil background, and combines multi-feature information in order to complete the automatic recognition of weeds in wheat fields [15], [16]. However, the features of wheat and weeds obtained through spectral imaging technology are still limited to two-dimensional space, which cannot fundamentally solve the issue of grass weed recognition and spectral imaging equipment is too expensive to truly apply in practical agricultural production.

Existing recognition methods for weeds in wheat fields mostly focus on wheat seedlings, but as the wheat grows, the leaf layers close the rows, and the degree of leaf overlap increases. The use of two-dimensional features greatly reduces the recognition effect of weeds in wheat fields. Studies have shown that plant height information is an important characterization of crop growth status, and weeds and crops often exhibit large differences in height due to growth competition and other factors [17], [18]. By acquiring the height information of crops and weeds, and combining image features and height features to classify weeds and crops, the accuracy of weed recognition is improved [19]. At present, the main technical methods for obtaining depth information include RGB-D cameras, lidar, and stereovision [20]. Although highly accurate height information can be obtained using radar, the equipment structure is complex and the cost is high, making it difficult to put into practical use in the field. When obtaining three-dimensional information via stereovision, the image-matching algorithm requires a large amount of calculation, which is not conducive to the real-time acquisition of weed depth in wheat fields. Therefore, developing a feasible method for obtaining the depth information of weeds in wheat fields quickly and at low cost is the first issue that needs to be solved in practical applications.

In order to overcome the inherent challenges in the recognition of weeds in wheat fields, we accomplished the following in this study:

1) A fast recognition method for weeds in wheat fields combining RGB images and depth images was proposed. For the first time, RGB-D fusion information was applied to the classification of weeds in wheat fields.

2) Focusing on the issue of holes in depth images, a depth information repair method based on RGB image information guidance was proposed that utilizes the object consistency of RGB and depth image acquisition.

3) A more robust classification algorithm of the various weed species in wheat fields was proposed. The different features of weeds in wheat fields at the peak of weed emergence were analyzed. Color, position, texture, and depth information was extracted. The AdaBoost algorithm was employed for integrated learning in order to achieve accurate recognition of weeds in wheat fields.

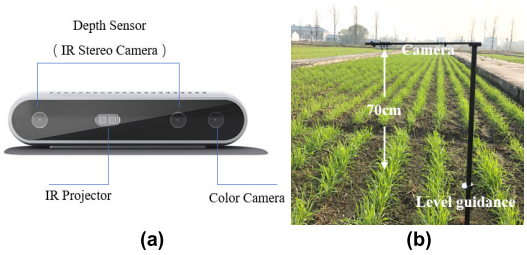


FIGURE 1. Image collection scene and acquisition equipment: (a) Intel®RealSense™D415 camera; (b) Image acquisition equipment.

In the follow sections, the experimental design and data acquisition methods will be introduced in section II. In section III, the distribution features of weeds in wheat fields at different stages are analyzed, and a data gap repair algorithm and classification method for weeds in wheat fields are proposed. Section IV provides the experimental results. Our conclusions are presented in the final section.

II. RGB-D DATASET COLLECTION

A. EXPERIMENTAL DESIGN

From December 2017–April 2018, wheat trials were conducted at the National Engineering and Technology Center for Information Agriculture Base in Rugao City, Nantong City, Jiangsu Province, China. The wheat varieties selected for the test were Shengxuan 6 (V1) and Sumai 8 (V2). There were 3 nitrogen treatments in the experiment, N_0 (0 kg N/ha), N_1 (180 kg N/ha), and N_2 (360 kg N/ha); and the test area was 100 m long and 12 m wide. Weeding was not carried out during field management. In addition, the seeds of 6 weed varieties commonly found in wheat fields were randomly sown in order to simulate the growth of field weeds. These included the grass weeds foxtail grass, annual bluegrass, brome, and barnyard grass, and the broad-leaved weeds redroot pigweed and shepherd's purse, a mix that was comparable to the actual range of weed varieties in wheat fields.

B. ACQUISITION EQUIPMENT

Wheat field RGB and depth images were acquired using the Intel® RealSense™Depth Camera D415 as shown in Fig. 1. The infrared stereo cameras were used to generate the depth images, and the color sensor was used to generate the RGB images, both with resolutions of 1280 × 720 pixels. Field images were obtained during the tillering and jointing stages of wheat, one being an RGB image and the other being a depth image. The camera was positioned 70 cm from the crop canopy. The images taken were based on a row of crops and soil on both sides. The actual field shooting device is shown in Fig. 1.

FIGURE 2. Distribution of weeds in wheat fields at different stages, the weeds framed in red: (a) Image of weeds during the tillering stage; (b) Image of weeds during the jointing stage(R2C2).

III. RECOGNITION METHODS FOR WEEDS IN WHEAT FIELDS

A. ANALYSIS OF DISTRIBUTION FEATURES OF WEEDS IN WHEAT FIELDS

There are 2 weed maxima during the growth of winter wheat. The first peak occurs approximately 40 days after the wheat is sown, specifically, before and after the tillering stage, and the second peak occurs around the jointing stage of the next year [8]. There are 2 types of weeds in wheat fields: between-row weeds and in-row weeds. During the tillering stage of wheat, the plants are low, there is no interconnection across crop rows, the spacing between rows is clear, weeds and wheat leaves basically do not overlap, and weeds between rows are easy to recognize, as shown in Fig. 2. (a). In order to analyze the distribution features of between-row weeds and in-row weeds in the tillering stage, we collected 50 RGB images during this wheat growth stage and manually counted the emergence of between-row weeds and in-row weeds. The statistical results, listed in TABLE 1, revealed that the total weed density is effectively determined by the density of between-row weeds at the tillering stage.

With the jointing and row-closing of wheat, the crop leaves become overlapped and blocked, and between-row weeds are no longer clearly visible. Weeds and wheat leaves overlap significantly, making it difficult to use spatial features to statistically classify the distribution of weeds, as shown in Fig. 2. (b). In the wheat field images, the proportion of soil pixels in the image decreases with the growth of wheat. Seeded wheat and clustered weeds usually present different textures. In addition, the height of wheat plants after jointing should also be different from the height of weeds. In order to prove this conjecture, 50 pixels of wheat and weeds were selected in the depth images of different experimental treatments in order to calculate the average height. The calculation results are shown in Fig. 3. These results demonstrated that for the 2 varieties and 3 different nitrogen treatments, wheat and weeds exhibited significant height differences, thus proving that height can be used as one of the classification features to distinguish wheat and weeds.

B. RECOGNITION METHOD FOR WEEDS IN WHEAT FIELDS DURING THE TILLERING STAGE

1) EXTRACTION OF THE FEATURES OF WEEDS IN WHEAT FIELDS IN THE TILLERING STAGE

During the tillering stage of wheat, the color of the plant contrasts distinctly with the soil, there are abundant between-row weeds, and the positions of wheat and weeds are significantly different. Therefore, we extracted position and color features for weed recognition. In terms of color features, the Excess Green (ExG) color vegetation index [21] was used to convert RGB images into grayscale images in order to remove the effect of soil pixels on subsequent segmentation. When calculating the ExG, the values of the 3 channels R, G, and B are first normalized, and the range of pixel values is reduced to [0,1]. A second normalization is then performed [22], so that the value of a given pixel in the 3 channels sums to 1. The calculation formula of the ExG index is as follows:

$$ExG = 2g - b - r \tag{1}$$

After using the color feature to complete the soil background segmentation, we employed the position feature to recognize the weeds located between the crop rows. The obtained gray image was binarized using Otsu's method of maximum class variance, thereby obtaining a binary image consisting of only wheat and weeds.

TABLE 1. Distribution of weeds in the tillering stage.

Wheat growth stage	Containing between-row weeds and in-row weeds	Containing in-row weeds but no between-row weeds	Containing between-row weeds but no in-row weeds
Tillering stage	6%	10%	84%

2) CLASSIFICATION AND EVALUATION OF WEEDS IN WHEAT FIELDS DURING THE TILLERING STAGE

Since the wheat images collected during the tillering stage contained only 1 crop row, we calculated the area of each connected domain in the binary image, and then removed the largest connected domain in the image, which was the crop row. The remaining connected domains were considered to be weeds. We then divided the obtained image into 225 64 $\times$ 64 image blocks, manually labeled the image blocks in each image, and compared and analyzed the results obtained by the algorithm in order to evaluate the accuracy of the classification method for weeds in wheat fields.

C. CLASSIFICATION METHOD FOR WEEDS IN WHEAT FIELDS DURING THE JOINTING STAGE

1) DEPTH IMAGE REPAIR

Although RealSense technology provides a low-cost, real-time way to obtain depth information, due to lighting conditions, infrared reflection characteristics of the surface material of the measured object, and area occlusion, there will be missing data, i.e., the issue of data gaps, which will decrease the accuracy of the target object information and affect subsequent feature extraction. We propose a data gap repair algorithm based on RGB information, including image alignment and gap repair.

Because the imaging origin of RGB images and depth images are not consistent, before performing data gap repair,

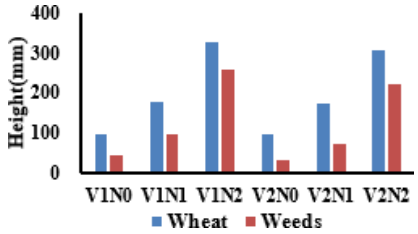


FIGURE 3. Results of height analysis of wheat and weeds under different treatments during the jointing stage(V1 and V2 represent test varieties; N0, N1, and N2 reflect nitrogen treatment.).

we aligned the coordinate systems of the depth images and the RGB images. The internal parameters matrix, rotation matrix, and translation vector of the depth camera and RGB camera were obtained from the System Design Kit (SDK) provided by Intel. The process of image alignment consists of first restoring the depth point of the depth image coordinate system to the world coordinate system, with P_{ir} set as the spatial coordinate of a point in the depth camera coordinate system, p_{ir} as the projection coordinate of the point on the image plane, and H_{ir} as the internal parameter matrix of the depth camera. As can be seen from the pinhole imaging model, these variables satisfy the following relationship:

$$p_{ir} = H_{ir}P_{ir} \tag{2}$$

$$P_{ir} = H_{ir}^{-1}p_{ir} \tag{3}$$

Next, the depth point in the world coordinate system is converted into the RGB image coordinate system, and P_{rgb} is set as the spatial coordinate of the same point in the RGB camera coordinate system, with p_{rgb} as the projection coordinate of the point on the RGB image plane, and H_{rgb} as the internal parameter matrix of the RGB camera. Since the coordinates of the depth camera and the RGB camera are different, they must be connected by rotation and translation transformation, where R is the rotation matrix and T is the translation vector. Specifically,

$$P_{rgb} = RP_{ir} + T \tag{4}$$

$$p_{rgb} = H_{rgb}P_{rgb} \tag{5}$$

The accuracy of the data gap repair has a great impact on the subsequent feature extraction. Since the object species of data gap points cannot be confirmed, the method of neighborhood interpolation may fill in data gap points with the depth information of other species, thereby causing an error [23]. The depth estimation algorithm requires a large amount of calculation, making it difficult to meet real-time requirements [24], [20]. In this section, we propose using the known object information in the RGB image as a priori knowledge and as a guide to fill data gaps.

After the depth information has been aligned with the RGB image, the data gap point index in the depth image will correspond to the pixel value of the RGB image pixel point. At the same time, the pixel value of the RGB image pixel point corresponding to the effective depth information in the 5×5 area centered on the data gap point is indexed. The pixel coordinates with the greatest similarity to the pixel points in the RGB image corresponding to the data gap points are then determined, and the gaps are repaired according to the depth information corresponding to the pixel coordinates. Equation (6) is the calculation formula of pixel similarity, where R_c, G_c, and B_c is the color channel information of the RGB image corresponding to the data gap point; R_i, G_i, and B_i is the color channel information of the RGB image corresponding to the effective point; $i = 1, \ldots N$; and N is the number of valid points within the 5×5 area.

$$S = |R_c - R_i| + |G_c - G_i| + |B_c - B_i| \tag{6}$$

2) EXTRACTING THE FEATURES OF WEEDS IN WHEAT FIELDS DURING THE JOINTING STAGE

Based on the analysis results of weed distribution during the jointing stage, texture and depth features were used to classify weeds in wheat fields. In the calculation process, segmenting a wheat field image into 64×64 image blocks to extract texture features effectively avoids the repetitive calculation of similar textures while considering regional consistency. Since the hue, saturation, intensity (HSI) color model can better express the hue and saturation information of the image, the HSI color space was used to calculate the texture information of wheat and weeds. We employed the color co-occurrence matrix (CCM) to extract texture features. The CCM not only considers spatial interaction between pixels, but also measures color distribution in the image [25]. The color co-occurrence matrix usually contains a large amount of information and cannot be used directly. In order to reduce the feature space dimensions as much as possible to decrease the amount of calculation, while still retaining the description of the image texture information, Maheswari selected 5 feature statistics—entropy, inverse moment, contrast, correlation, and angular second-order moment—to characterize texture [26].

The calculation process for the texture feature was as follows: First, the RGB image was subjected to color space conversion, and its corresponding value in the HSI color space model was calculated. In this study, the H, S, and I components were non-uniformly quantized into 8 parts in order to reduce color redundancy; this 8-level quantization has the added benefit of simplifying the calculation of the CCM. When calculating the CCM, the 512×512 image was divided into $64 \times 64 \times 64$ image blocks from left to right and top to bottom. The 3 components of the color space were set as C_1, C_2, and C_3, respectively; $m = C_k$ and $n = C_{k'}$ were the 2 components in the 3-color component combination space ($k, k' \in \{1, 2, 3\}$). Therefore, the color co-occurrence matrix $CCM_{m,n}$ was used to represent the measurement of the pixel color components C_k and $C_{k'}$ in the image; specifically, the space interaction between m and n. From this, the CCM_{HH}, CCM_{SS}, CCM_{II}, CCM_{HS}, CCM_{HI}, and CCM_{SI} matrices could be calculated. For any pixel in the image, assuming the value of the k^{th} color component is i, i.e., $m = i$

and the value of the k'^{th} color component is j, i.e., $n = j$, the elements $CCM_{m,n}(i, j)$ in the matrix are used to represent the number of times such pixels appear in the image. The formula for calculating the color co-occurrence matrix is as follows:

$$CCM_{m,n}(i,j) = \sum_{\Delta x}\sum_{\Delta y}\begin{cases} 1 & m(x+\Delta x, y+\Delta y) \\ & = i \& n(x+\Delta x, y+\Delta y) = j) \\ 0 & \text{others} \end{cases} \quad (7)$$

Here, the elements in the i^{th} row and j^{th} column of the color co-occurrence matrix are represented by $p(i, j)$, and L is the quantization level of the image. Each matrix calculates the 5 texture features of entropy, inverse moment, contrast, correlation, and angular second-order moment.

Although texture features can effectively recognize broad-leaved weeds in wheat fields, it is difficult to distinguish wheat from grass weeds, which affects the classification results. Therefore, the depth features of wheat and weeds were extracted from the depth information in order to recognize weeds in wheat fields during the jointing stage. Because the depth information displayed by a depth sensor is the straight line distance between the pixel and the sensor, there will be errors in the height of the pixels at the edge of the image, which is consistent with the RGB image. Similarly, a 512×512 area in the middle of depth image was selected for calculation.

The height information presented in depth images is the distance from the sensor to the target object. Therefore, during calculation, the maximum value, which is the height above the ground of the sensor, is first extracted, from which the depth information of each pixel is then subtracted in order to obtain the actual height of each pixel. Because the image block is small, the height information of each pixel in the region is similar. The depth information in the image block is de-zeroed and averaged to reflect the height features of all plants in the region. In addition, in the calculation of texture features, there are samples with a high proportion of soil pixels in the 64×64 image samples, which affects the classification of sample species and the calculation results of texture features. Therefore, we used the height information to remove samples in which the soil height (0 mm) pixels accounted for more than 90% of the total pixels.

3) CLASSIFICATION AND EVALUATION OF WEEDS IN WHEAT FIELDS DURING THE JOINTING STAGE

Due to the multi-dimensional features involved when using texture features to classify weeds in wheat fields, 3 types of classic classifiers were selected: the backpropagation (BP) neural network, k-nearest neighbors (k-NN), and support vector machine (SVM). The distribution of the dataset consisted of 192 pieces of 64×64 image blocks for the training set and test set of wheat, with a corresponding texture feature matrix size of 192×15. The training set and test set of weeds both consisted of 70 pieces of 64×64 image blocks, with a corresponding texture feature matrix size of 70×15. The depth features were extracted by comparing wheat canopy height

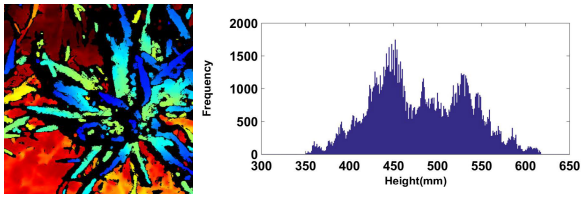


FIGURE 4. Histogram of depth information of the wheat field during the jointing stage: (a) Depth image; (b) Histogram of depth information.

with sample height information. Fig. 4. (b) is a histogram of the height information for all pixels in an image taken during the jointing stage of wheat. We defined the most frequent height information to be the canopy height of the wheat. The canopy height was then compared with the average value calculated in units of image blocks. Image blocks less than the canopy height were regarded as weeds, while image blocks greater than or equal to the canopy height were regarded as wheat. In order to combine texture features with depth features, Bagging (Bootstrap aggregating) and Boosting methods were used for ensemble learning respectively. The basic idea of both is to superimpose a large number of simple classifiers with general classification capabilities using a particular method in order to form a strong classifier with strong classification capabilities. In the Bagging algorithm [27], the training sets of each model are independent from each other and extracted from the original sample set, and the weight of each prediction function is equal. In Boosting algorithm, the weights of samples and classifiers will change and the AdaBoost algorithm [28] was used in this paper. First, the weight distribution of the data is initialized and the same weight is assigned to each sample. Then the weak classifier is trained, and the weights are assigned according to the accuracy of the classifier of weeds in wheat fields, which is based on texture features and depth features. The weight α_m calculation formula is (8), where m represents the number of iterations and e_m represents the error rate of the classifier.

$$\alpha_m = \frac{1}{2}\log\left(\frac{1-e_m}{e_m}\right) \quad (8)$$

The weight N update formula of samples is (9), where is the total number of samples, i represents the i^{th} sample, G represents the classifier, y is the sample label, and z_m is the normalization factor, such that the sum of the weights of all samples is 1.

$$w_{m+1,i} = \frac{w_{mi}}{z_m}\exp(-\alpha_m y_i G_m(x_i)) \quad (9)$$

$$z_m = \sum_{i=1}^{N} w_{mi}\exp(-\alpha_m y_i G_m(x_i)) \quad (10)$$

The evaluation method is divided into 2 steps. The first step is to evaluate the data gap repair algorithm. Due to the lack of a true value in the gap of the depth information of weeds collected in the field, it is difficult to evaluate the

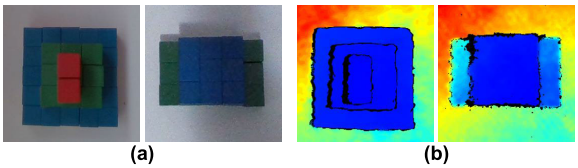


FIGURE 5. Test images: (a) RGB image; (b) Depth image.

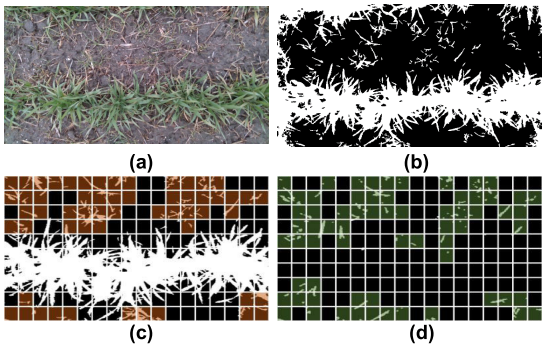


FIGURE 6. Weed recognition results during the tillering stage: (a) Original image of wheat tillering stage; (b) Binary image; (c) Manual labeling result; (d) Algorithm recognition result.

repair accuracy. We conducted tests on the building blocks (Fig. 5). The volume of each block was 25 × 25 × 25 mm and we used different color wooden blocks to build different structures. Since the true height of the block was known, the repair algorithm could be evaluated by comparing it with the repair results. We used RMSE to calculate the repair accuracy.

The second step is the evaluation of the classification method of weeds in wheat fields. The results of manual annotations were compared with the results of the algorithm calculations. In addition, the analysis of the weed species (broad-leaved and grass weeds) in the sample was performed in order to test the classification results of different methods for different weed species

IV. EXPERIMENTAL RESULTS

A. CLASSIFICATION RESULTS FOR WEEDS IN WHEAT FIELDS

The results of the algorithm during the tillering stage are shown in Fig. 6. By comparing the results of manual annotation of 50 photos with the recognition results of the algorithm, the rate of correct weed recognition was 88%, and the execution (run) speed of the algorithm was approximately 0.2 s. By analyzing the classification results of the algorithm for weeds in wheat fields during the tillering stage, it was discovered that a small portion of between-row weeds were close to the crop rows, and the areas of weeds and crop rows that were interconnected were thus difficult to distinguish.

B. CLASSIFICATION RESULTS FOR WEEDS IN WHEAT FIELDS DURING THE JOINTING STAGE

1) DATA GAP REPAIR ALGORITHM RESULTS

In order to reduce the loss of depth features and more accurately express the height information in the samples, the results of the data gap repair algorithm were analyzed.

TABLE 2. Data gap repair algorithm results.

Methods	RMSE	Run time
Our methods	2.02	0.37s

TABLE 3. SVM kernel function test results.

Kernel function	Classification accuracy
Polynomial function	80.15%
Gaussian kernel function	83.97%
Sigmoid function	78.24%

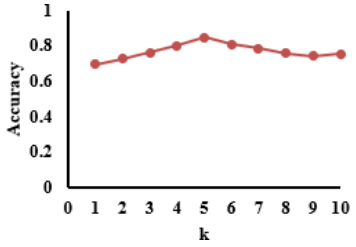


FIGURE 7. k-NN classification results for different k values.

TABLE 2 lists the calculation results of RMSE and the speed of the algorithm, proving that our algorithm can fill the missing information accurately and quickly.

2) CLASSIFIATION RESULTS FOR WEEDS IN WHEAT FIELDS BASED ON TEXTURE FEATURES

Before using wheat classifiers to classify weeds in wheat fields, we performed pre-tests to determine the structure and parameters of each classifier. During the structural design of the BP neural network, given the requirements for real-time performance, we designed 2 hidden layers. The transfer function of the first hidden layer is a hyperbolic tangent sigmoid function, and the transfer function of the second hidden layer is a logarithmic sigmoid transfer function. The training function of back propagation adopts the momentum gradient descent algorithm of variable learning rate. And the learning rate is set to 0.01. For the selection of the kernel function of the SVM, we calculated and compared the classification results of 3 kernel functions, i.e., the polynomial function, Gaussian kernel function, and sigmoid function. In Polynomial kernel function, the degree of the polynomial is 3. In Gaussian kernel function and sigmoid function, the kernel parameters are set to 0.5, this parameter is the inverse of the number of categories. TABLE 3 shows that the Gaussian kernel function performed well. For the setting of the hyperparameter k in the k-NN classifier, we also compared the classification effect of k ranging from 1–10, which revealed (Fig. 7) that the classification accuracy was relatively high when $k = 5$.

Considering that wheat varieties may affect the classification results of wheat and weeds, we selected 15 images under the same nitrogen treatment in each variety for texture calculation. Figure 9 shows the texture analysis results of two

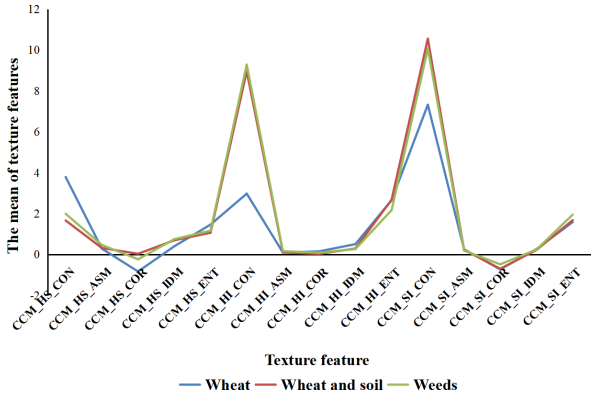


FIGURE 8. Texture feature analysis based on CCM_HS_ENT.

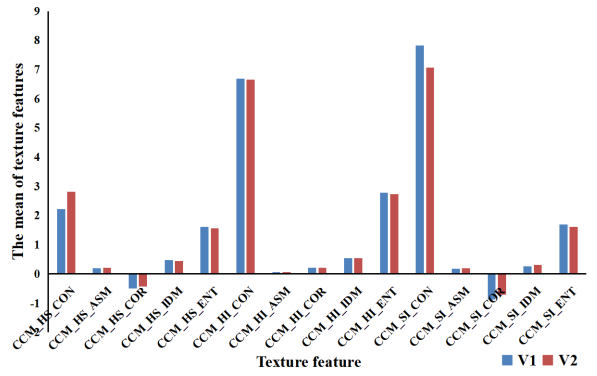


FIGURE 9. Texture feature analysis of two varieties of wheat.

wheat varieties and proves that the texture difference between different varieties is small. TABLE 4 shows the accuracy of the 3 classifiers based on texture features. It can be seen from the results that the recognition accuracy of wheat is significantly higher than that of weeds. After analyzing the sample set, we found that one reason for the poor accuracy of texture features in weed recognition was the impact of grass weeds on recognition accuracy. In addition, the image blocks at the edges of crop rows usually contain some soil, which is very similar to the content of the weed image blocks. We compared the texture calculation results of these 15 samples with the texture calculation results of 15 randomly selected weed samples. Fig. 8 shows the high similarity in the texture results of the 2 types of samples. Because the number of wheat samples was larger than the number of weed samples, the misrecognition of the wheat samples exerted a great impact on the accuracy of weed recognition.

3) CLASSIFICATION RESULTS FOR WEEDS IN WHEAT FIELDS BASED ON DEPTH FEATURES

We performed a height comparison between crop row edge samples and weed samples, and found that the height difference was large, as shown in Fig. 10. The classification results listed in TABLE 5 indicate that the accuracy of weed recognition based on depth features is significantly higher than that based on texture features. Since depth features are not sensitive to weed species information, the recognition

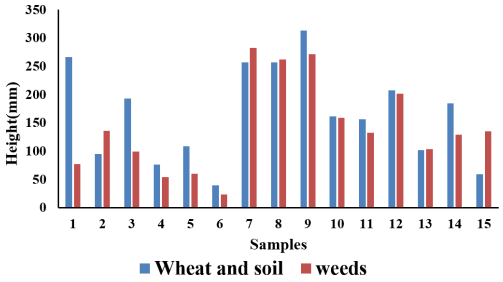


FIGURE 10. Depth feature analysis of samples.

FIGURE 11. Weed recognition results during the jointing stage.

TABLE 4. Classification results of weeds in wheat fields based on texture features from RGB images.

Class-ifier	Wheat	Weeds			Overall
		Grass weeds	Broad-leaved weeds	All weeds	
BP	91.15%	21.21%	48.65%	35.71%	76.34%
k-NN	95.31%	32.43%	78.79%	54.28%	84.73%
SVM	98.96%	24.32%	63.64%	42.86%	83.97%

TABLE 5. Classification results of weeds in wheat fields based on height information from depth images.

Feature	Wheat	Weeds			Overall
		Grass weeds	Broad-leaved weeds	All weeds	
Mean	86%	70.27%	72.72%	71.42%	84.73%

accuracy of grass weeds is also high. The recognition accuracy of wheat is slightly lower than the recognition results based on texture features, which may be due to the uneven distribution of the wheat canopy height.

4) INTEGRATED LEARNING RESULTS

The number of our data set is 524, including 384 wheat samples and 140 weed samples. In order to facilitate data set partitioning, 4-fold cross-validation is applied to integrated learning, with each data set containing 96 wheat samples and 35 weed samples. In the classification results of weeds in wheat fields based on texture features, the k-NN and SVM classifiers exhibited relatively high accuracy, so we employed the AdaBoost algorithm to calculate the weight distribution of the 3 classifiers (k-NN, SVM, and depth-based linear weed recognition), which were 0.34, 0.18, and 0.53, respectively. TABLE 6 shows that the identification accuracy of wheat based on Bagging is higher than the Ada Boost. Since all classifiers are equal in weight, the accuracy of wheat based on RGB image classifier is higher. However, the accuracy

TABLE 6. Classification results of weeds in wheat fields based on the fusion of RGB images and depth images.

Classifier	Wheat	Weeds			Overall	Run time
		Grass weeds	Broad-leaved weeds	All weeds		
Bagging	95.57%	42.19%	75.00%	60.26%	86.07%	0.54s
Ada Boost	88.80%	81.25%	81.58%	80.81%	86.19%	0.63s

of weeds based on Ada Boost is significantly higher than Bagging. Because in Ada Boost, the classifier based on depth information, which is sensitive for weed recognition, is given biggest weight.

The results demonstrated that the fusion of texture and depth features can recognize weeds in wheat fields well, and due to the addition of depth features, the algorithm is robust to weed species, and displays good recognition results for grass weeds and broad-leaved weeds. Because Ada Boost performs better in weeds recognition, we input the photos collected in the field into this model for the classification of weeds in wheat fields. Fig. 11 shows that the model can recognize most weeds. However, from the recognition results, we found that some image blocks contained both wheat and weed targets, and the overlap of the leaves of the 2 targets was serious, resulting in missed weed detection. In addition, due to the real-time nature of the method, we utilized hard segmentation in the acquisition of the samples and ignored the impact of different areas of weed blocks on the segmentation of image blocks. As a result, although small areas of weeds appeared in some image blocks, they were missed.

V. CONCLUSIONS

Due to the limitations inherent in the extraction of two-dimensional spatial features using RGB images, current research on the classification and recognition of weeds in wheat fields has mostly focused on the wheat tillering stage. Given the high similarity in appearance between grass weeds and wheat, the number of recognized weed species is limited, making it impractical for field application. In our proposed method, the distribution features of weeds at the peak of emergence in wheat fields were first analyzed in combination with actual agricultural production, and the classification features applied at different growth stages were extracted. Second, the depth feature of weeds in wheat fields extracted from depth information breaks through the limitation of two-dimensional spatial features in grass weed recognition. In addition, when performing depth feature extraction, we proposed a data gap repair method for the depth information in order to avoid feature loss caused by missing information. The test results revealed that the recognition accuracy levels of the depth features and texture features of broad-leaved weeds were similar, and the depth feature was 2.2 times the texture feature in the recognition accuracy of grass weeds. Using the AdaBoost algorithm to perform integrated learning on 3 outstanding classifiers (k-NN, SVM, and depth-based linear classifiers), the recognition accuracy of grass weeds was improved to 81.08%, which is 2.5 times the accuracy based on texture feature recognition.

The classification results of weeds in wheat fields during different growth stages revealed that the accuracy of weed recognition was 88% during the tillering stage of wheat, and it took approximately 0.2 s to calculate a 1280 × 720 image. During the jointing stage of wheat, the classification accuracy of weeds in wheat fields was 87.4%. Calculating a 512 × 512 image took about 0.69 s, which meets the needs of fast and accurate weeding in the field. In order to further improve the classification accuracy of weeds in wheat fields, the issue of missed weed detection due to the hard segmentation of images should be examined in future work. In addition, we will further study and calculate the volume and density of weeds based on the test results, so that they can be better applied to precise weeding in the field.

REFERENCES

[1] C. Spitters and J. Van Den Bergh, "Competition between crop and weeds: A system approach," in *Biology and Ecology of Weeds*. Dordrecht, The Netherlands: Springer, 1982, pp. 137–148.

[2] N. M. Munier-Jolain, S. H. M. Guyot, and N. Colbach, "A 3D model for light interception in heterogeneous crop:Weed canopies: Model structure and evaluation," *Ecol. Model.*, vol. 250, pp. 101–110, Feb. 2013.

[3] R. L. Zimdahl, *Fundamentals of Weed Science*. New York, NY, USA: Academic, 2018.

[4] A. R. Kniss, "Long-term trends in the intensity and relative toxicity of herbicide use," *Nature Commun.*, vol. 8, no. 1, p. 14865, Apr. 2017.

[5] J. R.-V. Osten and R. Dzul-Caamal, "Glyphosate residues in groundwater, drinking water and urine of subsistence farmers from intensive agriculture localities: A survey in Hopelchén, Campeche, Mexico," *Int. J. Environ. Res. Public Health*, vol. 14, no. 6, p. 595, Jun. 2017.

[6] M. T. Rose, T. R. Cavagnaro, C. A. Scanlan, T. J. Rose, T. Vancov, S. Kimber, I. R. Kennedy, R. S. Kookana, and L. Van Zwieten "Impact of herbicides on soil biology and function," in *Advances in Agronomy*, vol. 136. Amsterdam, The Netherlands: Elsevier, 2016, pp. 133–220.

[7] S. Christensen, T. Heisel, A. M. Walter, and E. Graglia, "A decision algorithm for patch spraying," *Weed Res.*, vol. 43, no. 4, pp. 276–284, Aug. 2003.

[8] S. Gaba, B. Chauvel, F. Dessaint, V. Bretagnolle, and S. Petit, "Weed species richness in winter wheat increases with landscape heterogeneity," *Agricult., Ecosyst. Environ.*, vol. 138, nos. 3–4, pp. 318–323, Aug. 2010.

[9] L. Ulber, H.-H. Steinmann, S. Klimek, and J. Isselstein, "An on-farm approach to investigate the impact of diversified crop rotations on weed species richness and composition in winter wheat," *Weed Res.*, vol. 49, no. 5, pp. 534–543, Oct. 2009.

[10] A. Farooq, J. Hu, and X. Jia, "Analysis of spectral bands and spatial resolutions for weed classification via deep convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 183–187, Feb. 2019.

[11] A. Piron, V. Leemans, O. Kleynen, F. Lebeau, and M.-F. Destain, "Selection of the most efficient wavelength bands for discriminating weeds from crop," *Comput. Electron. Agricult.*, vol. 62, no. 2, pp. 141–148, Jul. 2008.

[12] U. Shapira, I. Herrmann, A. Karnieli, and D. J. Bonfil, "Field spectroscopy for weed detection in wheat and chickpea fields," *Int. J. Remote Sens.*, vol. 34, no. 17, pp. 6094–6108, Sep. 2013.

[13] A. T. Nieuwenhuizen, J. W. Hofstee, and E. J. van Henten, "Performance evaluation of an automated detection and control system for volunteer potatoes in sugar beet fields," *Biosyst. Eng.*, vol. 107, no. 1, pp. 46–53, Sep. 2010.

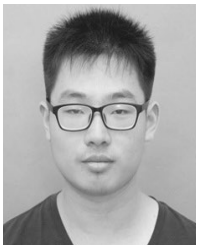
[14] A. Tellaeche, G. Pajares, X. P. Burgos-Artizzu, and A. Ribeiro, "A computer vision approach for weeds identification through support vector machines," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 908–915, Jan. 2011.

[15] J.-J.-Q. Armstrong, R. D. Dirks, and K. D. Gibson, "The use of early season multispectral images for weed detection in corn," *Weed Technol.*, vol. 21, no. 4, pp. 857–862, Dec. 2007.

[16] I. Sa, Z. Chen, M. Popovic, R. Khanna, F. Liebisch, J. Nieto, and R. Siegwart, "WeedNet: Dense semantic weed classification using multispectral images and MAV for smart farming," *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 588–595, Jan. 2018.
[17] A. Piron, V. Leemans, F. Lebeau, and M.-F. Destain, "Improving in-row weed detection in multispectral stereoscopic images," *Comput. Electron. Agricult.*, vol. 69, no. 1, pp. 73–79, Nov. 2009.
[18] L. Zhang and T. E. Grift, "A LiDAR-based crop height measurement system for miscanthus giganteus," *Comput. Electron. Agricult.*, vol. 85, pp. 70–76, Jul. 2012.
[19] A. Piron, F. van der Heijden, and M. F. Destain, "Weed detection in 3D images," *Precis. Agricult.*, vol. 12, no. 5, pp. 607–622, Oct. 2011.
[20] C. Yan, B. Gong, Y. Wei, and Y. Gao, "Deep multi-view enhancement hashing for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Feb. 24, 2020, doi: 10.1109/TPAMI.2020.2975798.
[21] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multi-view RGB-D object dataset," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2011, pp. 1817–1824.
[22] V. Surazhsky, M. Bronstein, A. Bronstein, R. Kimmel, E. Sperling, A. Zabatani, O. Menashe, and D. H. Silver, "Code filters for coded light depth acquisition in depth images," U.S. Patent 9 792 671 B2, Oct. 17, 2017.
[23] G. E. Meyer and J. C. Neto, "Verification of color vegetation indices for automated crop imaging applications," *Comput. Electron. Agricult.*, vol. 63, no. 2, pp. 282–293, Oct. 2008.
[24] C. Yan, B. Shao, H. Zhao, R. Ning, Y. Zhang, and F. Xu, "3D room layout estimation from a single RGB image," *IEEE Trans. Multimedia*, early access, Jan. 17, 2020, doi: 10.1109/TMM.2020.2967645.
[25] C. Palm, "Color texture classification by integrative co-occurrence matrices," *Pattern Recognit.*, vol. 37, no. 5, pp. 965–976, May 2004.
[26] G. U. Maheswari, K. Ramar, D. Manimegalai, and V. Gomathi, "An adaptive region based color texture segmentation using fuzzified distance metric," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2916–2924, Mar. 2011.
[27] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.
[28] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

KE XU received the B.E degree in electronic engineering from Anhui Polytechnic University, Wuhu, China, in 2017. She is currently pursuing the Ph.D. degree in agricultural informatics with Nanjing Agricultural University, Nanjing, China.

Since September 2017, she has been engaged in research on computer vision and machine learning with the National Information Agricultural Engineering Technology Center, China.



HUAIMIN LI received the B.Sc. (Agriculture) degree in agronomy from Nanjing Agricultural University, Nanjing, China, in 2019, where he is currently pursuing the M.Sc. (Agriculture) degree in agricultural informatics.

His research interests include crop growth spectral monitoring and image processing.

WEIXING CAO received the Ph.D. degree in crop science from Oregon State University, in 1989.

He is currently a Professor with Nanjing Agricultural University, where he has been engaged in research on crop ecology and information agriculture. He is also the Dean of the Intelligent Agricultural Research Institute, Nanjing Agricultural University, and the Leader of the National Agricultural Professional Degree Collaboration Group, Agricultural Engineering and Information Technology, China.

YAN ZHU received the Ph.D. degree in agricultural informatics from Nanjing Agricultural University, in 2013.

She is currently a Professor with Nanjing Agricultural University, where she has been engaged in research on crop modeling and information agriculture. She is also the Dean of the College of Agricultural, Nanjing Agricultural University, and the Director of the National Information Agricultural Engineering Technology Center.

RONGJIA CHEN is currently pursuing the bachelor's degree in agronomy with Nanjing Agricultural University, Nanjing, China.

Her research interests include agricultural informatics and intelligent agriculture.

JUN NI received the Ph.D. degree from the Institute of Agricultural Engineering, Jiangsu University, Zhenjiang, China.

He is currently a Professor with the National Information Agricultural Engineering Technology Center; the Engineering Research Center of Smart Agriculture? Ministry of Education; the Key Laboratory for Crop System Analysis and Decision Making, Ministry of Agriculture and Rural Affairs; and the Jiangsu Key Laboratory for Information Agriculture, the Jiangsu Collaborative Innovation Center for the Technology and Application of Internet of Things, Nanjing Agricultural University, China. His research interests include crop information perception and processing and agricultural sensor and intelligent equipment, in which he has published 36 journal articles, filed 25 patents, and 13 computer software copyrights.

Dr. Ni is a member of the Artificial Intelligence Branch of the Chinese Society for Agricultural Machinery and the Agricultural Knowledge Engineering Committee of the Chinese Association of Automation.