



CNN feature based graph convolutional network for weed and crop recognition in smart farming



Honghua Jiang^{a,1}, Chuanyin Zhang^{a,1}, Yongliang Qiao^{b,*}, Zhao Zhang^c, Wenjing Zhang^a, Changqing Song^a

^a College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China

^b Australian Centre for Field Robotics (ACFR), Faculty of Engineering, The University of Sydney, NSW 2006, Australia

^c College of Electronic and Electrical Engineering, Baoji University of Arts and Sciences, Baoji 721016, China

ARTICLE INFO

Keywords:

Weed recognition
Graph convolutional network
Deep learning
Semi-supervised learning
Smart farming

ABSTRACT

Weeding is an effective way to increase crop yields. Reliable and accurate weed recognition is a prerequisite for achieving high-precision site-specific weed control in precision agriculture. To improve weed and crop recognition accuracy, a CNN feature based graph convolutional network (GCN) based approach is proposed. A GCN graph was constructed based on extracted weed CNN features and their Euclidean distances. Based on the semi-supervised learning, the GCN graph enriched the model by exploiting labeled and unlabeled image features, and testing samples obtain label information from labeled weed data by performing propagation over the graph. The proposed GCN-ResNet-101 approach achieved 97.80%, 99.37%, 98.93% and 96.51% recognition accuracies on four different weed datasets respectively, which outperformed the state-of-the-art methods (AlexNet, VGG16 and ResNet-101). Additionally, the runtime of the proposed approach also satisfies the real-time requirement of field weed control. The proposed CNN feature based GCN approach is favorable for multi-class crops and weeds recognition with limited labeled data, which is a promising approach in dealing with similar agricultural recognition tasks. Furthermore, the used datasets and source code are publicly available to facilitate the research in the recognition of field weeds.

1. Introduction

Adequate agricultural products are prerequisites to sustain our worldwide population growth. One of the approaches to food supply with minimal or no environmental damage (Wang et al., 2019; Chavan and Nandedkar, 2018) is relying on new modern technologies (e.g. sensors, big data or deep learning) or so-called smart farming. Smart farming technology achieves yield increases by extracting key indicators of crop growth and conducting high-precision weed control strategies at different growth stages (Sa et al., 2018).

Weeds are regarded as one of the main threats in the output of agricultural production because they cause significant yield loss by competing with crops for nutrients, sunlight, space and water. Weeds also cause depreciation of product quality and host insects or diseases (Olsen et al., 2019). For weed control, spraying of herbicides is the most common and effective way. However, the traditional uniform herbicide spraying does not distinguish crops and weeds, which leads to excessive waste. In addition, uniform herbicide spraying causes great pollution to

the soil and water sources. Therefore, there is a growing need to recognize weeds and crops automatically with high accuracy for site-specific herbicide applications.

In early stages, human-craft features such as the color (Zheng et al., 2017), texture (Dongjian et al., 2013) and shape (Bakhshipour and Jafari, 2018) are widely used for weed detection and recognition. Nowadays, inspired by the successful performance of Convolutional Neural Networks (CNN) (Yang et al., 2019; Qiao et al., 2019), many researchers proposed CNN based weed classification or recognition approaches (Bah et al., 2018; Chavan and Nandedkar, 2018). Bah et al. (2018) proposed a CNN based automatic weed detection and recognition approach using UAV images. Sa et al. (2018) proposed dense semantic weed recognition using multi-spectral images collected by a micro aerial vehicle (MAV). Chavan and Nandedkar (2018) proposed AgroAVNET (a hybrid model of AlexNet and VGGNET) for crop-weed species classification. Farooq et al. (2019) used the CNN method to extract more discriminative and powerful features that result in accurate weed classification. However, the CNN based deep learning

* Corresponding author.

E-mail address: yongliang.qiao@sydney.edu.au (Y. Qiao).

¹ The Authors contributed equally and share first authorship.

approach requires a very large amount of manual labeled images for model training. Unfortunately, in terms of field weed recognition, that is not always possible due to extremely time-consuming data labeling.

To improve weed and crop recognition accuracy with limited labeling data, one potential approach is semi-supervised learning, where a large amount of unlabeled data can be leveraged in training with typically a small amount of labeled data to significantly improve learning accuracy (Zhu et al., 2019). The recently emerged GCN is a successful semi-supervised learning approach, which has already achieved state-of-the-art performance in the tasks of text classification, image and speech recognition (Wang et al., 2019). GCN learns new representations of a vertex by aggregating feature vectors of all neighbours in the aggregation process (Kipf and Welling, 2016). Based on semi-supervised learning, GCN can significantly improve recognition accuracy with small amounts of labeled data.

In this work, in order to improve weed and crop recognition accuracy with limited labels in challenging field environments, a CNN feature based GCN is proposed by taking advantage of CNN features and the semi-supervised learning ability of the graph. Firstly, all weed (training and testing samples) CNN features are extracted, then the GCN graph is constructed based on weed features (graph vertex) and their Euclidean distances (graph edge). In the GCN model, the graph performs feature propagation between labeled and unlabeled graph vertices through graph convolution layers. During this process, unlabeled vertices (testing samples) can update their features, and the final obtained features are fed into the softmax classifier of labeled vertices for label prediction. Thus unlabeled weed samples are recognized based on the proposed approach.

The main contributions of this work are: (1) Introducing a GCN based semi-supervised learning approach for weed recognition with a small number of labeled data, which take advantage of CNN features and the feature propagation ability of graph; (2) Results of four different datasets demonstrate the proposed approach outperformed the other state-of-the-art methods—AlexNet, VGG16 and ResNet-101; (3) Weed recognition benchmark datasets, containing 6000 images of four different weed species and corn seedlings have been made publicly available. The used experimental datasets and source code are available at <https://github.com/zhangchuanyin/weed-datasets>.

The rest of the paper is organized as follows: Section 2 briefly reviews some recent weed recognition methods, Section 3 introduces the used datasets and the proposed approach, Section 4 presents all experimental results, Section 5 discusses the application's specifications, and then conclusions and future areas for research are given in Section 6.

2. Related works

Automatic weed and crop recognition using computer vision algorithms is an important academic and practical challenge (Bakhshipour and Jafari, 2018). Aimed at recognizing weeds accurately, extensive studies and research have been conducted.

2.1. Feature-based approaches

The early weed and crop recognition approaches mainly utilize human-craft features such as color (Zheng et al., 2017), texture (Dongjian et al., 2013) and shape (Bakhshipour and Jafari, 2018).

In one study (Zheng et al., 2017), color indices and support vector data description were used to distinguish maize from different species of weeds under various weather conditions. In terms of shape and texture features, Dongjian et al. (2013) fused shape, texture and fractal dimensions as weed features, then support vector machine (SVM) and DS (Shafer-Dempster) evidence theory were used for weed recognition. Bakhshipour and Jafari (2018) integrated shape factors, Fourier descriptors, and moment invariant features to establish a crop-weed discrimination pattern, and the proposed approach achieved 92.92%

overall classification accuracy. In addition, Kounalakis et al. (2018) proposed an image-based weed recognition system using SURF features and an optimized Gaussian Mixture Model. The proposed framework was found to be efficient for using weed control robots for precision farming applications with a recognition accuracy of 89.09%. In addition, spectral features have also been adopted in weed recognition (Sa et al., 2018). Lottes et al. (2017) proposed crop and weed recognition method for smart farming using RGB and near infra-red (NIR) imagery taken from the UAV.

However, most of the above mentioned artificially-designed features rely on human experience and have a certain degree of blindness, and recognition performance varies greatly between different datasets (Bah et al., 2018).

2.2. Deep learning based approaches

Compared with the traditional human-craft features (e.g. color, texture and shape), the weed features extracted by powerful CNN are more robust, and they are always more representative and invariant to the illumination or soil background (Dyrmann et al., 2017; Tang et al., 2017). Dyrmann et al. (2017) trained and validated a CNN model to recognize weeds from winter wheat fields with heavy leaf occlusion. The model yielded a recall of 46.3% and a precision of 86.6%. Tang et al. (2017) proposed a k-means unsupervised feature learning approach to replace the random initialization weights of traditional CNN parameters, and the proposed approach improve weed identification accuracy.

More recently, Olsen et al. (2019) used Inception-v3 and ResNet-50 deep learning models to achieve an average weed classification accuracy of 95.1% and 95.7% respectively on their own multi-class weed image dataset. Xiangwu et al. (2018) proposed a multi-feature fusion and deep confidence network for rice seedling weed identification, and the proposed approach achieved a 91.13% recognition rate. A unsupervised deep learning and semi-automatic data labeling approach for weed discrimination was proposed in work (dos Santos Ferreira et al., 2019), which achieved 97% accuracy in discrimination of grass and broadleaf weeds while reducing the number of manual annotations by 100 times.

However, the majority of current deep learning approaches are only focused on specific datasets without considering the influence of illumination and soil background (Tang et al., 2017). In addition, these deep learning approaches only achieve acceptable recognition accuracy based on large scale labeled data training, which is not suitable to deal with the agricultural recognition tasks that only have limited labeled data (Wang et al., 2019). How to construct an accurate weed and crop recognition model with strong learning ability on limited label data still is a challenge.

3. Material and methods

3.1. Weed datasets

For weed recognition experiments, four different datasets are used, namely corn, lettuce, radish, and mixed weed datasets. The corn and lettuce weed datasets are acquired by ourself, while the third one is from the public radish weed dataset (Lameski et al., 2017). The mixed dataset was constructed by combining corn, lettuce and radish weed datasets together. All the datasets used are open to the public².

Our corn weed dataset was taken from the natural environment of the corn seedlings field in 2016. A Canon PowerShot SX600 HS camera was used vertically towards the ground to acquire images that can reduce the influence of the sunlight reflection. As displayed in Fig. 1, four different weed species (i.e. Cirsium setosum, Chenopodium album,

² <https://github.com/zhangchuanyin/weed-datasets>.

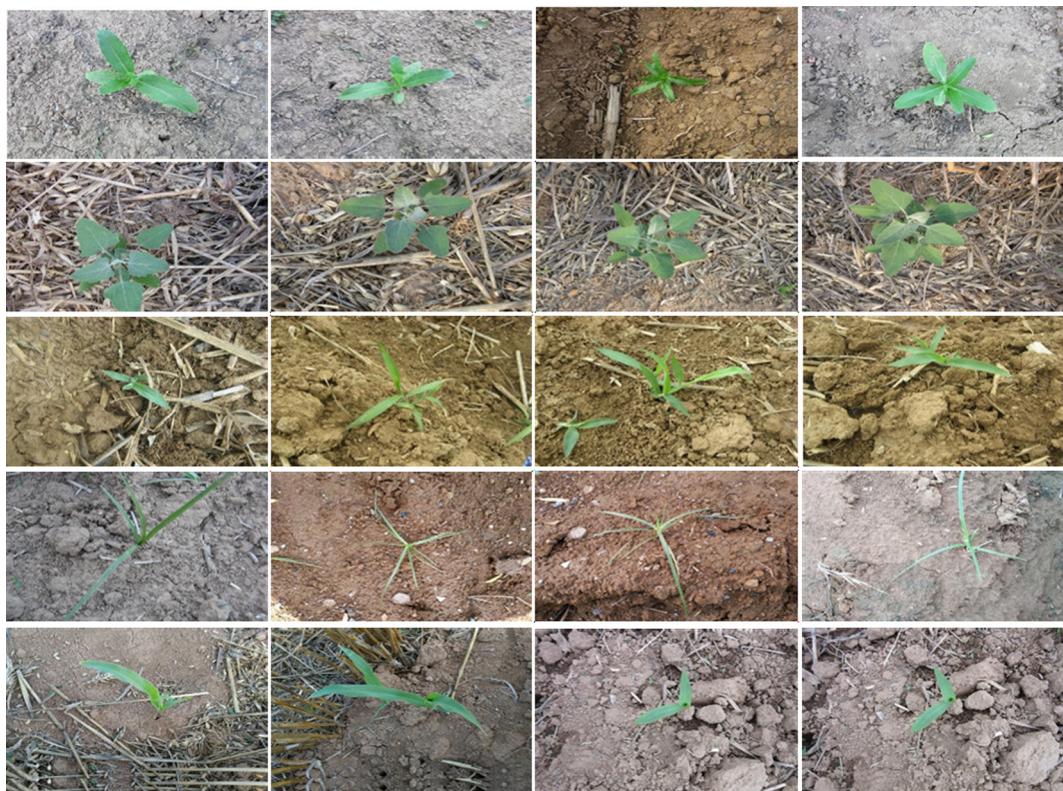


Fig. 1. Example images from the corn weed dataset. The first four rows are weeds, and each row is one species of weed which is *Cirsium setosum*, *Chenopodium album*, bluegrass and sedge respectively. The last row is corn.

bluegrass and sedge) and corn images were collected on different days, illumination and soil background (e.g. moisture and wheat straw residue) situations. A total of 6000 weed and corn images were collected. In the experiment, original images are resized to 800×600 .

The example images of lettuce and radish weed dataset are shown in Fig. 2. Images were collected from a vegetable plantation. The data collection was from two different days, and in total 500 lettuce seedling images and 300 weed images were acquired from a 30 cm height. The public radish weed dataset contains 39 RGB images taken from an approximately 1 m height (Lameski et al., 2017). Considering the number of images is small and the size of image is large, we cropped the original large images (3264×2448) into smaller images (800×600). Finally, 200 images of radish seedlings and 200 images of its associated weeds are obtained and used in our experiments.

3.2. Overview of the proposed approach

Accurately recognizing weeds is a prerequisite for variable-rate spraying requirement in precision agriculture farming (Tang et al., 2017; Olsen et al., 2019). The proposed CNN feature based GCN method mixed the features of unlabeled vertices with those of nearby labeled vertices, thus the problem of weed and crop recognition is transferred to semi-supervised learning on a graph.

As illustrated in Fig. 3, the proposed method consists of two parts: (1) CNN feature extraction: each image feature is extracted using a CNN model, and we can finally obtain a CNN feature set, where $N = N_{train} + N_{test}$ is the total number of images, N_{train} and N_{test} being the numbers of training and testing images respectively, C is extracted feature dimension; (2) GCN recognition model: a GCN graph $G = (X, V)$ is constructed based on extracted CNN feature X and vertex set V . Here V consists of V_{train} and V_{test} , where V_{train} and V_{test} are labeled and unlabeled vertices separately. The goal of semi-supervised learning with



Fig. 2. Example images from the lettuce and radish weed datasets. Images in the upper row are from the lettuce weed dataset, while images at the bottom row are from the radish weed dataset.

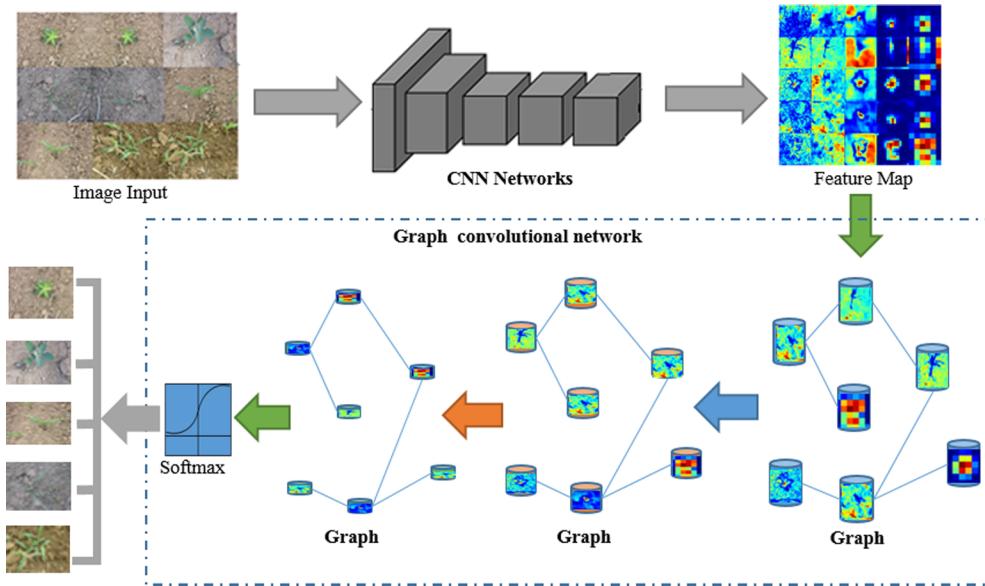


Fig. 3. The proposed CNN feature based GCN for the recognition of weed and crop.

GCN is to obtain the label information of unlabeled vertices V_{test} based on labeled vertices V_{train} .

The GCN model naturally combines graph structures and vertex features in the convolution, and propagated over the graph through multiple layers. Through the graph convolution layer by layer, unlabeled vertices V_{test} update their features. Then the final obtained features of unlabeled vertices are fed into the softmax classifier, and the output is label information of unlabeled vertices V_{test} . Thus all unlabeled images are recognized as weed or crop based on the proposed approach. Here the softmax classifier is trained based on features of labeled vertices V_{train} .

3.3. CNN feature extraction

CNN networks are used to extract CNN features due to their strong feature extraction ability (Krizhevsky et al., 2012; Qiao et al., 2019). In CNN networks, original images were input, then the convolution operation of each layer was performed to acquire image features. The lower layer contains more spatial details while the higher layer has more semantic information (Qiao et al., 2019). In order to obtain robust CNN features, the output of the final pooling layer is used as the image features.

In order to compare the performance of different CNN features, three different kinds of CNN features are extracted from AlexNet (Krizhevsky et al., 2012), VGG16 (Simonyan and Zisserman, 2014) and ResNet-101 (He et al., 2016) respectively. Each network is firstly fine-tuned with our dataset, and then fully connected layers are removed and only the front part is retained for weed feature extraction. Each image feature is extracted using the CNN model, and the final obtained CNN feature set X will be used to construct the GCN graph.

3.4. GCN based weed and crop recognition model

3.4.1. Graph construction

To construct GCN graph G , two elements are needed: (1) CNN feature set X ; (2) Matrix of feature relationship $\bar{A}: \bar{A} \in R^{N \times N}$ is a Laplacian transformed adjacency matrix, where $\bar{A} = \widetilde{D}^{-1/2}(A + I_N)\widetilde{D}^{-1/2}$ (D is a diagonal matrix, A is an adjacency matrix, I_N is an N -dimensional unit matrix).

In this work, the adjacency matrix A is the Euclidean distance of each two weed image features. The high value in the adjacency matrix A means that the weeds have high similarity.

Following the strategies in Kipf and Welling (2016), a two-layer GCN (not including the input layer) is used to recognize weed and crop. After GCN graph G is constructed, the graph vertex features are weighted and fused to the adjacent vertices based on adjacency matrix A . The neighboring nodes acquire similar features.

3.4.2. Graph convolution for GCN

GCN generalizes traditional convolutional neural networks to the graph domain. Spatial GCN and Spectral GCN are mainly two types of graph convolutional networks. Considering the real-time requirement for weed identifying in real crop planting, a type of spectral convolution is used in this work. The spectral convolution of CNN feature set X on convolution kernel g_θ can be approximated by Chebyshev polynomial (Kipf and Welling, 2016):

$$g_\theta \star X \approx \theta(I_N + D^{-1/2}AD^{-1/2})X \quad (1)$$

where \star denotes the convolution operator.

GCN as an effective graph representation model that can naturally combine structure information and vertex features in the learning process, which represents a vertex by aggregating feature vectors from its neighbors (including itself) (Zhu et al., 2019). The propagation rule of GCN can be summarized by the following expression:

$$H^{(l+1)} = \theta(\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}H^{(l)}W^{(l)}) \quad (2)$$

where $H^{(l)}$ is the matrix after convolutional operation in the l -th layer, $H^{(0)} = X$, $\widetilde{A} = A + I_N$, $\widetilde{D} = \sum_j \widetilde{A}_{ij}$, $W^{(l)}$ is the trainable weight matrix in the l -th layer, and θ is the activation function.

3.4.3. Semi-supervised learning for weed and crop recognition

Based on two-layer graph convolution processing, unlabeled vertices V_{test} update their features. The final obtained features of unlabeled vertices are fed into the softmax classifier. Here the softmax classifier is trained based on labeled vertices V_{train} . The outputs of softmax classifier are labels (weed or crop) of unlabeled vertices V_{test} . Thus all unlabeled images are recognized based on the proposed approach. The whole process of the proposed two-layer GCN model can be formulated as:

$$Z = f(X, \widetilde{A}) = \text{Softmax}(\widehat{A} \text{Relu}(\widehat{A} X W^{(0)}) W^{(1)}) \quad (3)$$

where $\widehat{A} = \widetilde{D}^{-1/2}(A + I_N)\widetilde{D}^{-1/2}$.

The loss function is defined as the crossentropy error over all labeled vertices:

Table 1
Description of the datasets employed in the experiments.

Dataset	No.images	Description
Corn	Train: 4200 (840 corn and 3360 weed) Test: 1800 (360 corn and 1440 weed)	Major variations in illumination and soil background
Lettuce	Train: 560 (350 lettuce and 210 weed) Test: 240 (150 lettuce and 90 weed)	Major variations in illumination
Radish	Train: 280 (140 radish and 140 weed) Test: 120 (60 radish and 60 weed)	Minor variations in illumination
Mixed	Train: 5040 (840 corn, 350 lettuce, 140 radish and 3710 weed) Test: 2160 (360 corn, 150 lettuce, 60 radish and 1590 weed)	Multiclass crops minor variations in illumination

Table 2
Training parameters for networks.

Parameters	Value	Remarks
Batch size	32	Compute the gradient in a reasonable speed
Training epochs	20,000	Make sure the network is well trained
Learning rate	0.00001	Learn a more optimal set of weights
Initial weights	Random	No influence of pre-trained weights

Table 3
Performance comparison of different methods on the corn weed dataset (%).

Methods	ACC	Precision	Recall	SPC	F_1	Time (s)
AlexNet	93	97.9	97.76	91.98	97.83	0.18
VGG16	95.8	98.63	98.92	94.7	98.78	0.21
ResNet-101	96.5	98.8	98.83	95.3	98.81	0.28
GCN-AlexNet (ours)	95.6	98.71	98.57	95.01	98.64	1.03
GCN-VGG16 (ours)	97.2	99.06	99.15	96.3	99.10	1.12
GCN-ResNet-101 (ours)	97.8	99.26	99.23	97.1	99.25	1.24

$$L = \sum_{i \in Y_D} \sum_{f=1}^F Y_{if} \ln Z_{if} \quad (4)$$

where Y_D is the set of indices of labeled vertices and F is equal to the number of classes.

4. Experimental setup

4.1. Experimental dataset construction and network training

To validate the proposed GCN based weed and crop recognition, experiments were conducted on three different weed datasets (i.e. corn,

lettuce and radish) and a mixed dataset. The mixed dataset was constructed by combining corn, lettuce and radish weed datasets together. The details about the used datasets are summarized in [Table 1](#).

- (1) Corn weed dataset: 4200 images (i.e. 3360 weed and 840 corn seedling images) were used for GCN model training, and the remaining 1800 images (i.e. 1440 weed and 360 corn seedling images) were used for testing.
- (2) Lettuce weed dataset: 560 images (i.e. 350 lettuce seedling and 210 weed images) and 240 images (i.e. 150 lettuce seedling images and 90 weed images) were used to train and test the network respectively.
- (3) Radish weed dataset: 280 images (i.e. 140 radish seedling and 140 its associated weed images) were used for model training, while 120 images (i.e. 60 radish seedling and 60 weed images) were used for testing.
- (4) Mixed dataset: 5040 images (i.e. 840 corn, 350 lettuce, 140 radish and 3710 weed) were used for model training, while 2160 images (i.e. 360 corn, 150 lettuce, 60 radish and 1590 weed) were used for testing.

4.2. Comparison with the state-of-the-art methods

In addition, the proposed GCN based approaches were also compared with the state-of-the-art CNN models: AlexNet ([Krizhevsky et al., 2012](#)), VGG16 ([Simonyan and Zisserman, 2014](#)) and ResNet-101 ([He et al., 2016](#)). AlexNet consists of five convolutional layers and three fully-connected layers. VGG16 is a pretty large network with about 138 million (approx) parameters, but only 16 layers that have weights. ResNet-101 is released by Microsoft Research Asia, it contains 101 layers by learning the residual representation functions instead of learning the signal representation directly. These three models have achieved significant performance in the classification tasks, therefore they were selected and compared with our approaches.

4.3. Network training parameters

In this work, all experiments were conducted on a computer equipped with a GeForce GTX 1080 Ti GPU, I9-7920X CPU@2.9 GHz. In the following experiments, for a fair comparison between different networks, the parameters of batch size, learning rate, training epochs and initial weights for these networks are the same. All the networks' input size is $224 \times 224 \times 3$. As presented in [Table 2](#), for the network training on each dataset, the training epoch is set to 20,000, batch size is 32, and the learning rate is 0.00001. In addition, in order to avoid the influence of pre-trained weights, all the network's initial weight is

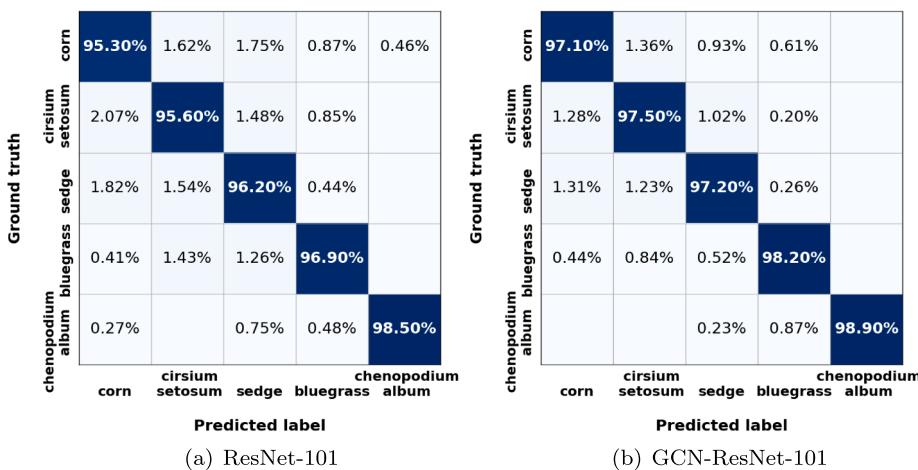


Fig. 4. Confusion matrix of the corn weed dataset recognition.



Fig. 5. Recognition results of the proposed GCN-ResNet-101 in corn weed dataset.

Table 4
Performance comparison of different methods on the lettuce weed dataset (%).

Methods	ACC	Precision	Recall	SPC	F_1	Time (s)
AlexNet	96.65	93.59	97.33	96	95.42	0.16
VGG16	97.61	95.77	98	97.4	96.87	0.19
ResNet-101	98.25	96.73	98.83	98	97.52	0.20
GCN-AlexNet (ours)	98.23	97.73	98.83	98	97.52	0.2
GCN-VGG16 (ours)	98.82	97.70	99	98.6	98.34	0.28
GCN-ResNet-101 (ours)	99.37	98.36	99.67	99	99	0.31

random. The other parameters of each network are their default settings.

4.4. Performance evaluation

The results of testing for all approaches were arranged in confusion matrices, including true positive (tp), true negative (tn), false positive (fp), and false negative (fn). In this context, tp represents the weeds that are correctly identified weeds; tn represents the crops that are correctly identified as crops; fp represents the crops that are incorrectly identified as weeds; and fn represents weeds are incorrectly identified as crops.

In order to evaluate performance, five widely used measures were

calculated: accuracy (ACC), precision, recall, specificity (SPC) and F_1 score. Accuracy is the ratio of number of correct predictions to the total number of input samples; precision shows the ability of the model to accurately identify targets; recall reflects the ability of the model to detect targets; F_1 score is a harmonic means of the precision and recall; specificity measures the proportion of actual negatives that are correctly identified. All the above five measures are ranged from 0 to 1, high value means the good predictive ability of the model, their definitions are as follows:

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \quad (5)$$

$$\text{Precision} = \frac{tp}{tp + fp} \quad (6)$$

$$\text{Recall} = \frac{tp}{tp + fn} \quad (7)$$

$$\text{Specificity} = \frac{tn}{tn + fp} \quad (8)$$

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

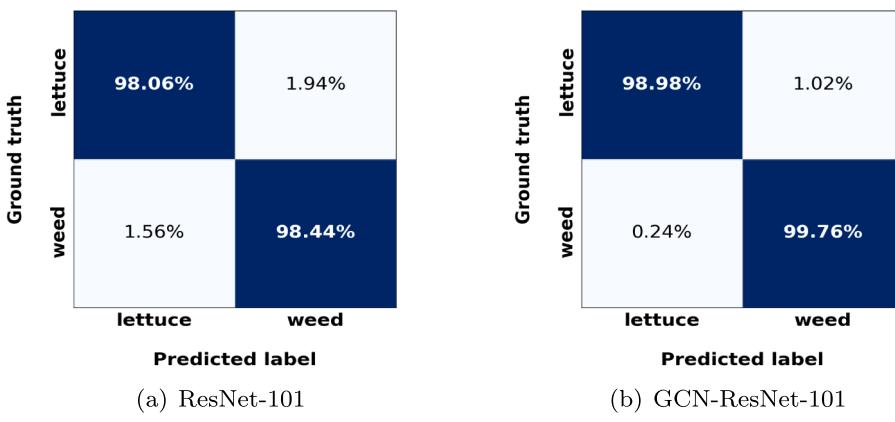


Fig. 6. Confusion matrix of the lettuce weed dataset recognition.

5. Experimental results

5.1. Corn weed dataset results

The recognition accuracy and runtime of CNN and GCN based approaches for corn weed dataset are illustrated in [Table 3](#). The proposed GCN-ResNet-101 based approach achieved an accuracy of 97.8%, a precision of 99.26%, a recall of 99.23%, a specificity of 97.1%, and an F_1 score of 99.25%. These yielded values are slightly higher than that of GCN-AlexNet and GCN-VGG16. It also can be noticed that the specificity of the proposed GCN based approach is almost 2% higher than their corresponding CNN models—AlexNet, VGG16 and ResNet-101. The results illustrate that the proposed GCN based approach could enhance the recognition ability through semi-supervised learning.

Table 5
Performance comparison of different methods on radish weed dataset (%).

Methods	ACC	Precision	Recall	SPC	F_1	Time (s)
AlexNet	95.85	95	96.7	95	95.89	0.15
VGG16	97.54	97.24	97.86	97.22	97.55	0.18
ResNet-101	97.95	97.83	98.08	97.82	97.95	0.21
GCN-AlexNet (ours)	97.50	98.27	96.7	98.3	97.48	0.23
GCN-VGG16 (ours)	98.42	98.81	98.12	98.82	98.46	0.30
GCN-ResNet-101 (ours)	98.93	98.54	9.33	98.53	98.93	0.32

In Fig. 4, confusion matrix of the corn weed dataset recognition is presented. It can be seen that the proposed GCN-ResNet-101 achieved 97.10% and 97.50% accuracies for corn and *Cirsium setosum*

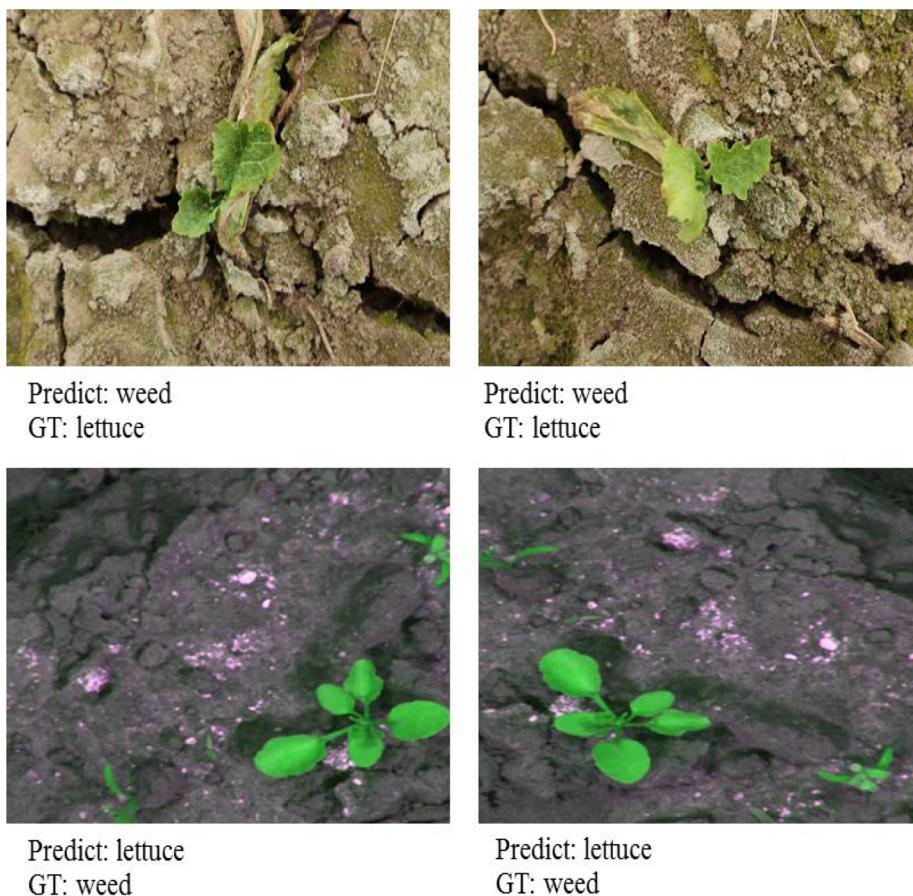


Fig. 7. False recognition examples of GCN-ResNet-101 approach in the lettuce weed dataset.

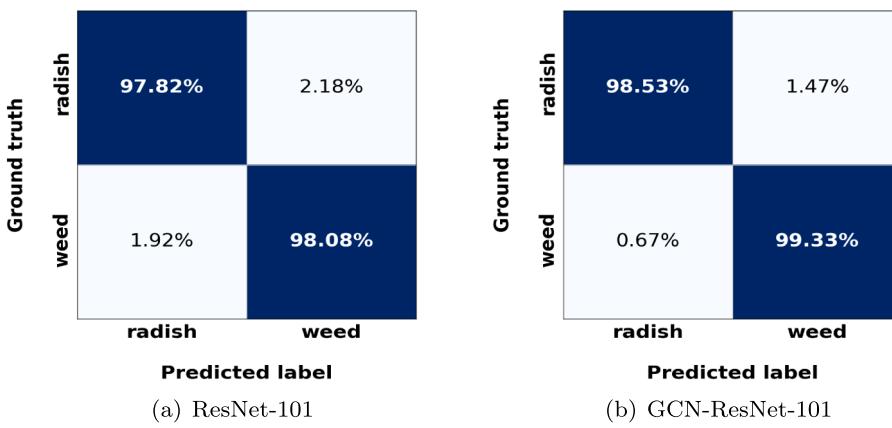


Fig. 8. Confusion matrix of the radish weed dataset recognition.

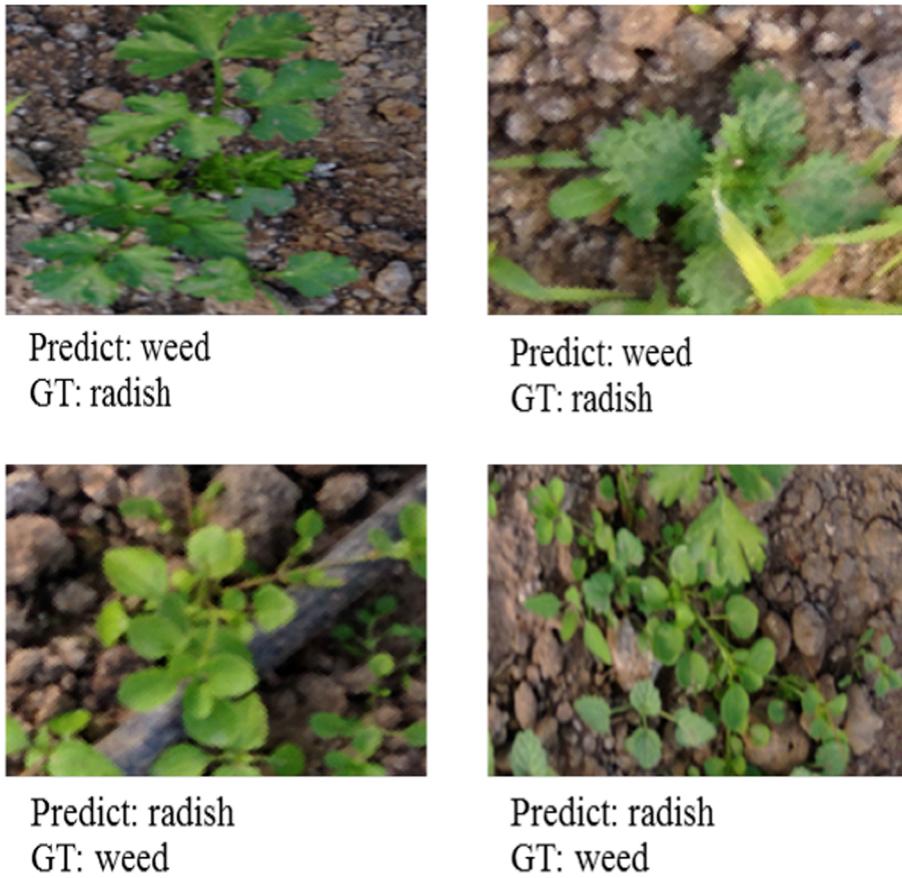


Fig. 9. False recognition examples of GCN-ResNet-101 approach in the radish weed dataset.

Table 6

Performance comparison of different methods on the mixed dataset (%).

Methods	ACC	Precision	Recall	SPC	F ₁	Time (s)
AlexNet	90.3	95.65	95.53	90.17	95.58	0.19
VGG16	91.37	96.03	95.98	91.03	96.05	0.23
ResNet-101	94.95	98.37	97.58	94.35	968	0.29
GCN-AlexNet (ours)	92.46	96.24	96.22	91.95	96.11	1.08
GCN-VGG16 (ours)	93.29	97.27	96.84	92.88	96.36	1.15
GCN-ResNet-101 (ours)	96.51	98.83	98.73	96.48	97.18	1.42

respectively, which is higher than that of the ResNet-101 (95.30% for corn and 95.60% for Cirsium setosum). In terms of weed species such as sedge and bluegrass, the recognition accuracy of the proposed GCN-

ResNet-101 is also higher than the ResNet-101. This is because the GCN exploits the relationship of CNN features which improved the distinguishing ability between different species.

		Predicted label			
		corn	lettuce	radish	weed
Ground truth	corn	94.17%	0.28%	0.56%	5%
	lettuce	0.67%	95.33%	0.67%	3.33%
	radish	1.67%		95%	3.33%
	weed	2.96%	1.13%	0.63%	95.28%

		Predicted label			
		corn	lettuce	radish	weed
Ground truth	corn	96.11%			3.89%
	lettuce		97.33%	0.67%	2%
	radish	1.67%		96.67%	1.67%
	weed	2.58%	1.01%	0.5%	95.91%

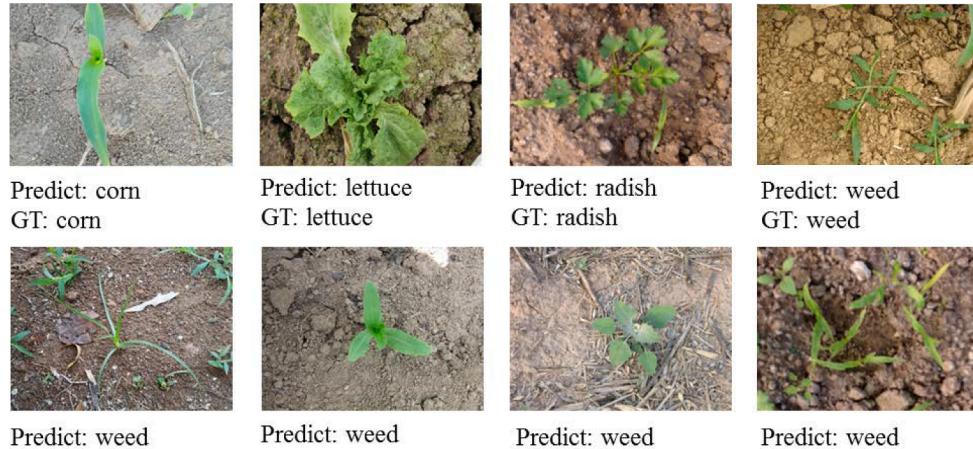
(a) ResNet-101

(b) GCN-ResNet-101

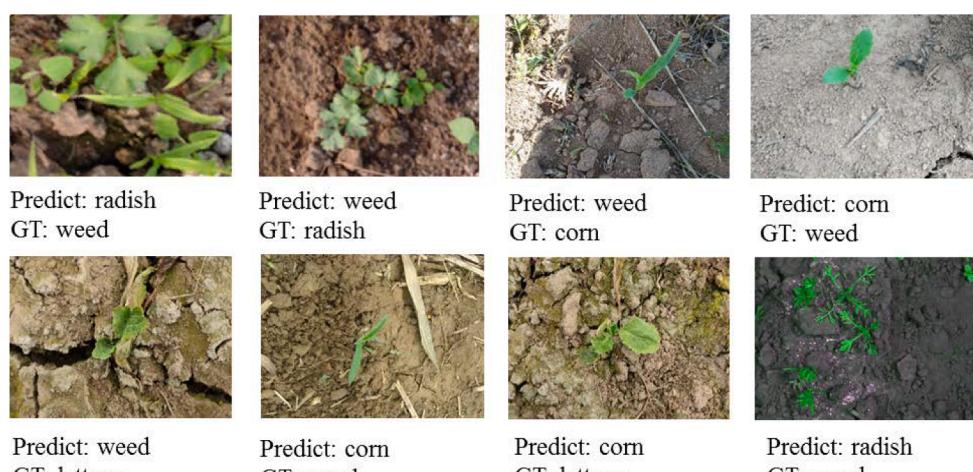
Fig. 10. Confusion matrix of the crop and weed recognition on mixed dataset.

In addition, some correct and false recognition examples of the proposed GCN-ResNet-101 are also shown in Fig. 5. It can be seen that there are a small number of corn samples that were incorrectly

recognized as weed. One possible reason is that the large area of the soil background in image influences the model learning and feature extraction during the network training. It is believed that small seedling



(a) Correct recognition examples



(b) False recognition examples

Fig. 11. Recognition results of the proposed GCN-ResNet-101 in the mixed dataset.

images increased in the training dataset can reduce this kind of false recognition.

5.2. Lettuce weed dataset results

In [Table 4](#), recognition accuracies of the lettuce weed dataset are illustrated. The proposed GCN-ResNet-101 achieved 99.37% recognition accuracy which is higher than that of GCN-AlexNet (98.23%), GCN-VGG16 (98.82%), ResNet-101 (98.25%), VGG16 (97.61%) and AlexNet (96.65%). Also it can be seen that F_1 score of proposed GCN-ResNet-101 (99%) is higher than that of the ResNet-101 (97.52%).

In addition, the confusion matrix of the lettuce weed dataset recognition is displayed in [Fig. 6](#). It can be seen that the proposed GCN-ResNet-101 achieved 99.76% weed recognition accuracy, which is higher than that of ResNet-101 (98.44%). The experimental results illustrate again that the proposed GCN-ResNet-101 has a strong feature representation ability for weed recognition which outperformed the AlexNet, VGG16 and ResNet-101 methods.

Some false recognition examples of proposed GCN-ResNet-101 are also shown in [Fig. 7](#). The possible reasons for false recognition are because of the different soil background and similar leaf shapes between the lettuce and weed seedlings. More variation in the dataset being considered in model training could be helpful to enhance the recognition accuracy.

5.3. Radish weed dataset results

The confusion matrix of the radish weed dataset recognition are illustrated in [Table 5](#). According to [Table 5](#), the recognition accuracy of the proposed GCN-ResNet-101 is 98.93%, which is higher than that of the ResNet-101 (97.95%), VGG16 (97.54%) and AlexNet (95.85%).

In [Fig. 8](#), the proposed GCN-ResNet-101 achieved 98.53% and 99.33% accuracies for radish and weed recognition respectively, which is higher than that of ResNet-101. The results illustrate again that the GCN based approach can exploit feature relationship and structure information to improve the recognition performance.

Additionally, [Fig. 9](#) displays some false recognition examples of the proposed GCN-ResNet-101 in the radish weed dataset. The situation of frequent leaf overlapping is the main reason for the false identification.

5.4. Performance on mixed dataset

The recognition accuracy of the proposed approach on mixed dataset are illustrated in [Table 6](#). The proposed GCN-ResNet-101 based approach achieved an accuracy of 96.51%, a precision of 98.83%, a recall of 98.73%, a specificity of 96.48%, and an F_1 score of 97.18%. These yielded values are higher than that of GCN-VGG16 and GCN-AlexNet. It also can be found that the performance of our GCN based approaches is almost 2% higher than their corresponding CNN models—AlexNet, VGG16 and ResNet-101. That means the proposed GCN based approach learns new feature representation from the neighbor nodes through graph convolution, which improves the recognition ability based on semi-supervised learning. Furthermore, the runtime of different approaches for testing dataset was also reported, and although the GCN based approaches took a little more time than the CNN, it is worth considering because of the improved performance.

In addition, [Fig. 10](#) shows the confusion matrix of the crop and weed recognition on mixed dataset. It can be seen that the proposed GCN-ResNet-101 for weed recognition is 95.91%, which is higher than that of ResNet-101 (95.28%). It can be found that most of the crops (i.e. corn, radish and lettuce) are correctly recognized, and only a small number of samples are incorrectly recognized as weeds. The results illustrate that our proposed approach is favorable for multi-class crops and weeds recognition in precision weed control of smart farming.

Some correct and false recognition examples of the proposed GCN-ResNet-101 are shown in [Fig. 11](#). It can be seen that some young lettuce

seedlings in extremely dry soil are falsely recognized as corn or weed. One explanation for this is that the soil moisture influences the crop growth, and some not well-growing crops are prone to be mis-recognized due to their abnormal leaf morphology (e.g. tiny or yellow leaves).

6. Discussion

From the results, we can see the proposed GCN-ResNet-101 approach outperforms AlexNet, VGG16 and ResNet-101, which is favorable for weed and crop recognition. The main reasons why the GCN based weed recognition works well are twofold: (1) the graph can capture feature relationships; (2) in the process of GCN semi-supervised learning, as a special form of Laplacian smoothing, the new features of each node are learned. The label information of weed sample vertices can be passed to their neighbouring vertices, and each vertex can gather comprehensive label information and act as bridges or key paths in the graph, so that label information can be propagated to the entire graph.

However, we also observed that there are some false recognition cases (see [Figs. 5\(b\), 7, 9 and 11\(b\)](#)) in the results of the proposed GCN approach. One reason is that some weeds and crops are very similar at their young seedling stage (corn vs Cirsium setosum). Another possible reason is that the GCN used here is not very deep as our dataset is not very large. Increasing the layer number and training images could further improve recognition accuracy. In addition, the adjacency matrix construction method in the GCN model also has an influence on the final recognition performance. The currently used Euclidean distance based adjacency matrix computes distances of each two vertices in the graph, which is a little low-efficiency. Actually, not all the vertices have significant contribution, and a focus on the K-nearest neighbour vertices would be more effective.

7. Conclusions

In order to improve the recognition accuracy of crop and weed with a limited labeled dataset, a semi-supervised GCN-ResNet-101 network is proposed. The proposed GCN-ResNet-101 consists of two parts: CNN feature extraction and GCN based recognition. The extracted robust CNN features are good representations of weed and corn images whilst semi-supervised GCN exploiting the graph or manifold structure of data and learning with limited labels. Combining the advantages of CNN features and a semi-supervised learning ability of the graph, the proposed GCN-ResNet-101 approach improves the recognition accuracy of crops and weeds by exploiting feature relationships and maximizing the effective utilization of the limited labeled data. It is a step towards precision weeding in smart agriculture.

In addition, the proposed approach was compared with the state-of-the-art methods (AlexNet, VGG16 and ResNet-101) on four different datasets. The experimental results show that the proposed GCN-ResNet-101 approach achieved 97.80%, 99.37%, 98.93% and 96.51% recognition accuracy on four different datasets respectively, which outperformed state-of-the-art methods—AlexNet, VGG16 and ResNet-101. Furthermore, the runtime of the proposed GCN method is 1.24s for 1800 testing image frames in corn and weed dataset, which satisfied the real-time requirement of field weed control. The dataset and source code is open to the public and it is hoped helpful to further research into the recognition of field weeds.

In the future, more efficient graph construction techniques will be explored and evaluated on an image dataset covering a greater range of variables (e.g. different locations, soils and image acquisition heights).

CRediT authorship contribution statement

Honghua Jiang: Writing - original draft, Resources, Funding acquisition. **Chuanyin Zhang:** Writing - original draft. **Yongliang Qiao:** Conceptualization, Methodology, Formal analysis, Writing - review &

editing. Zhao Zhang: **Wenjing Zhang**: Writing - review & editing.
Changqing Song: Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the support of Major Scientific and Technological Innovation Project of Shandong Province (2019JZZY010716), Major Agricultural Application Technology Innovation Project of Shandong Province (SD2019NJ001), Shandong Key R&D Program (2015GNC112004), National Key R&D Program (2017YFD0700500) and Shandong Natural Science Fund Project (ZR2018MC017).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.compag.2020.105450>.

References

- Bah, M., Hafiane, A., Canals, R., 2018. Deep learning with unsupervised data labeling for weed detection in line crops in uav images. *Remote Sens.* 10, 1690.
- Bakhshipour, A., Jafari, A., 2018. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* 145, 153–160.
- Chavan, T.R., Nandedkar, A.V., 2018. Agroavnet for crops and weeds classification: a step forward in automatic farming. *Comput. Electron. Agric.* 154, 361–372.
- Dongjian, H., Yongliang, Q., Pan, L., et al., 2013. Weed recognition based on svm-ds multi-feature fusion. *Trans. Chinese Soc. Agric. Mach.* 44, 182–187.
- dos Santos Ferreira, A., Freitas, D.M., da Silva, G.G., Pistori, H., Folhes, M.T., 2019. Unsupervised deep learning and semi-automatic data labeling in weed discrimination. *Comput. Electron. Agric.* 165, 104963.
- Dyrmann, M., Jørgensen, R.N., Midtiby, H.S., 2017. Robowedsupport-detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network. *Adv. Anim. Biosci.* 8, 842–847.
- Farooq, A., Hu, J., Jia, X., 2019. Analysis of spectral bands and spatial resolutions for weed classification via deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* 16, 183–187.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
- Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv:1609.02907.
- Kounalakis, T., Triantafyllidis, G.A., Nalpantidis, L., 2018. Image-based recognition framework for robotic weed control systems. *Multimedia Tools Appl.* 77, 9567–9594.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105.
- Lameski, P., Zdravevski, E., Trajkovik, V., Kulakov, A., 2017. Weed detection dataset with rgb images taken under variable light conditions. In: International Conference on ICT Innovations, Springer, pp. 112–119.
- Lottes, P., Khanna, R., Pfeifer, J., Siegwart, R., Stachniss, C., 2017. Uav-based crop and weed classification for smart farming. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 3024–3031.
- Olsen, A., Konovalov, D.A., Philippa, B., Ridd, P., Wood, J.C., Johns, J., Banks, W., Grgenti, B., Kenny, O., Whinney, J., et al., 2019. Deepweeds: A multiclass weed species image dataset for deep learning. *Sci. Rep.* 9, 2058.
- Qiao, Y., Truman, M., Sukkarieh, S., 2019. Cattle segmentation and contour extraction based on mask r-cnn for precision livestock farming. *Comput. Electron. Agric.* 165, 104958.
- Qiao, Y., Cappelle, C., Ruichek, Y., Yang, T., 2019. Convnet and LSH-based visual localization using localized sequence matching. *Sensors* 19, 2439.
- Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., Siegwart, R., 2018. weednet: Dense semantic weed classification using multispectral images and mav for smart farming. *IEEE Robot. Automat. Lett.* 3, 588–595.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.
- Tang, J., Wang, D., Zhang, Z., He, L., Xin, J., Xu, Y., 2017. Weed identification based on k-means feature learning combined with convolutional neural network. *Comput. Electron. Agric.* 135, 63–70.
- Wang, A., Zhang, W., Wei, X., 2019. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* 158, 226–240.
- Xiangwu, D., Long, Q., Xu, M., Yu, J., Xueshen, C., Haiyun, L., Weifeng, C., 2018. Recognition of weeds at seedling stage in paddy fields using multi-feature fusion and deep belief networks. *Trans. Chinese Soc. Agric. Eng.* 34, 165–172.
- Yang, T., Cappelle, C., Ruichek, Y., El Baghdouri, M., 2019. Multi-object tracking with discriminant correlation filter based deep learning tracker. *Integr. Comput.-Aided Eng.* 26, 273–284.
- Zheng, Y., Zhu, Q., Huang, M., Guo, Y., Qin, J., 2017. Maize and weed classification using color indices with support vector data description in outdoor fields. *Comput. Electron. Agric.* 141, 215–222.
- Zhu, R., Dornaika, F., Ruichek, Y., 2019. Joint graph based embedding and feature weighting for image classification. *Pattern Recogn.* 93, 458–469.