Crop Protection 174 (2023) 106389

Contents lists available at ScienceDirect

Crop Protection

journal homepage: www.elsevier.com/locate/cropro

Semi-supervised learning and attention mechanism for weed detection in wheat

Teng Liu [a], Xiaojun Jin [a,b], Luyao Zhang [a], Jie Wang [c], Yong Chen [b], Chengsong Hu [a,**], Jialin Yu [a,*]

[a] Peking University Institute of Advanced Agricultural Sciences/Shandong Laboratory of Advanced Agricultural Sciences at Weifang, Weifang, Shandong, China
[b] College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, Jiangsu, China
[c] Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao, 266101, China

ARTICLE INFO

Keywords:
Deep learning
Semi-supervised learning
Attention mechanism
Site-specific weed management

ABSTRACT

Machine vision-based precision herbicide application in wheat (*Triticum aestivum* L.) can substantially reduce herbicide input. However, detecting newly emerged weeds in wheat fields remains a challenging task. Current deep learning-based weed detection methods require the annotation of a large amount of data, which is both time-consuming and labor-intensive. To address this issue, this research improved a semi-supervised learning (SSL) algorithm based on consistency regularization and pseudo-labeling, and incorporated an attention mechanism. Compared to fully supervised learning (FSL) algorithms, the proposed method increased the classification accuracy by 16.5%, 17.84%, and 19.67% on datasets with 200 × 200, 300 × 300, and 400 × 400 pixel images, respectively, when only 100 labeled data per class were used. Overall, the developed machine vision models using the proposed method achieved weed detection with high accuracy while requiring much fewer labeled training images, and thus is more time and labor-efficient compared to an FSL algorithm.

1. Introduction

Wheat (*Triticum aestivum* L.) is an important agricultural crop widely cultivated in many parts of the world (Asseng et al., 2011). Wheat is one of the most important food sources and is a significant contributor to global food security (Igrejas and Branlard, 2020; Curtis and Halford, 2014). Weeds compete with wheat for resources, such as nutrients, water, and sunlight, thus hindering wheat growth. As a result, weed control is essential to ensure wheat productivity (Khan et al., 2011). In the past several decades, herbicide-resistant weed biotypes in wheat have been increasingly documented (Norsworthy et al., 2012; Peterson et al., 2018; Heap, 2023); however, broadcast-spraying synthetic herbicides is still the most extensively used strategy for weed control (Xiao et al., 2020; Wang et al., 2020). In natural conditions, weeds are randomly distributed in wheat fields, while broadcast application of herbicides leads to the application in areas where weeds do not occur. Manual spot spraying herbicides can reduce herbicide inputs but is impractical for large wheat fields (Bàrberi, 2002).

Accurate and reliable weed detection is the foundation for realizing autonomous spot-spraying herbicides. Previous research has utilized leaf feature differences to distinguish crops and weeds (Wu et al., 2021). For example, weed identification has been performed by extracting features such as plant color (Woebbecke et al., 1995; Tang et al., 2000; Meyer and Neto, 2008), leaf texture (Burks et al., 2000; Wu and Wen, 2009; Bakhshipour et al., 2017), shape (Kazmi et al., 2015; Bakhshipour and Jafari, 2018), and spectra (Franz et al., 1991; Shirzadifar et al., 2020) of weeds through image processing techniques. Some scholars have improved the accuracy of weed identification by extracting multiple features of weeds (Zhao et al., 2013; Lin et al., 2017; Torres-Sánchez et al., 2015). However, these methods are primarily based on manual design or existing statistical methods to extract features from images, which can be cumbersome and complex.

With the significant advancement in computing technology, traditional image processing techniques are being increasingly replaced by deep learning (O'Mahony et al., 2020). As a subset of machine learning technology, deep learning can perform complex feature extraction, efficiently process large quantities of data, and has demonstrated capabilities in various agricultural fields, such as crop classification

* Corresponding author.
** Corresponding author.
E-mail addresses: jialin.yu@pku-iaas.edu.cn (C. Hu), huchengsong0814@gmail.com (J. Yu).

https://doi.org/10.1016/j.cropro.2023.106389
Received 15 March 2023; Received in revised form 6 August 2023; Accepted 20 August 2023
Available online 24 August 2023

(Murmu and Biswas, 2015; Zhong et al., 2019), pest and disease recognition (Cheng et al., 2017; Chandy, 2019; Ahmad Loti et al., 2021), farmland management (Kamilaris and Prenafeta-Boldú, 2018), crop yield prediction (Khaki and Wang, 2019; Van Klompenburg et al., 2020; Nevavuori et al., 2019), and crop growth analysis (Bernotas et al., 2019; Yasrab et al., 2021). Deep learning convolutional neural networks (DCNNs) demonstrated excellent performances for detecting weeds growing in various cropping systems, such as corn (*Zea mays* L.) (Andrea et al., 2017; Ahmad et al., 2021), soybean [*Glycine* max (L.) Merr] (dos Santos Ferreira et al., 2017; Razfar et al., 2022), dormant or active-growing turfgrass (Yu et al., 2019a; Yu et al., 2019b; Yu et al., 2019c), and vegetables (Jin et al., 2021; Kennedy et al., 2020). Nevertheless, developing effective DCNN models for weed recognition requires collecting a large number of training images and performing image data labeling. For example, Zhuang et al. (2021) reported that the image classification neural networks trained with FSL using a very large training dataset containing 22,000 labeled training images effectively detected newly emerged broadleaf weed seedlings growing in wheat. However, Yan and Wang (2022) highlighted that obtaining sufficient labeled training data is often impossible or costly. In addition, some weed morphological features are similar to those of crops, thus requiring experienced weed scientists to perform the labeling tasks. Consequently, there is a need for an approach to training weed detection models with only a small amount of labeled data while preserving accurate detection.

SSL, a branch of deep learning, involves training on a small amount of labeled data and a large amount of unlabeled data. SSL can perform better with substantially less labeled data than supervised learning, which only uses labeled data (Van Engelen and Hoos, 2020). The core methods for semi-supervised image classification tasks are consistency regularization (Laine and Aila, 2016) and pseudo-labeling (Scudder, 1965), which are often combined in semi-supervised algorithms (Berthelot et al., 2019a,b; Sohn et al., 2020; Zhang et al., 2021). Consistency regularization is a method in semi-supervised learning aimed at maintaining the stability of a model to small perturbations in input data. This method is based on the assumption that the model's predictions should be consistent for similar input samples. Consistency regularization achieves this by introducing data augmentation techniques to apply random perturbations to input data and requiring the model to produce similar outputs for both the original and perturbed data. This improves the model's robustness and generalization ability on unlabeled data (Laine and Aila, 2016). Pseudo-labeling is used to generate artificial labels on unlabeled data. This method first trains the model using available labeled data and then applies the trained model to unlabeled data, generating pseudo-labels for the unlabeled data based on the model's predictions. This allows a portion of the unlabeled data to be treated as labeled data, further expanding the training dataset and improving the model's performance and generalization ability (Scudder, 1965). These unlabeled data can be transformed into labeled data with pseudo-labels by utilizing pseudo-labeling and consistency regularization methods. As a result, this method helps the model better utilize the information in the unlabeled data, thereby increasing the size of the training dataset, adding more training samples, and improving the accuracy and stability of the model. The confidence in these pseudo-labels is mostly high, regardless of whether the pseudo-labels are correct. However, if a large number of unlabeled samples are given incorrect labels and used for training, it can result in a large number of noisy samples in the training dataset, adversely affecting the performance of the model (Rizve et al., 2021).

In recent years, the scientific community has demonstrated growing research interest in the utilization of SSL algorithms to deal with various tasks such as text classification (Yin et al., 2015; Miyato et al., 2016), natural language processing (Søgaard, 2013), face recognition (Gao et al., 2015), building defect detection (Guo et al., 2021), and precision medicine (Chen et al., 2019; Wang et al., 2020). In agriculture, scholars recently used SSL to detect rows of crops to separate weeds from crops (Nong et al., 2022; Pérez-Ortiz et al., 2015) and assess weed density and

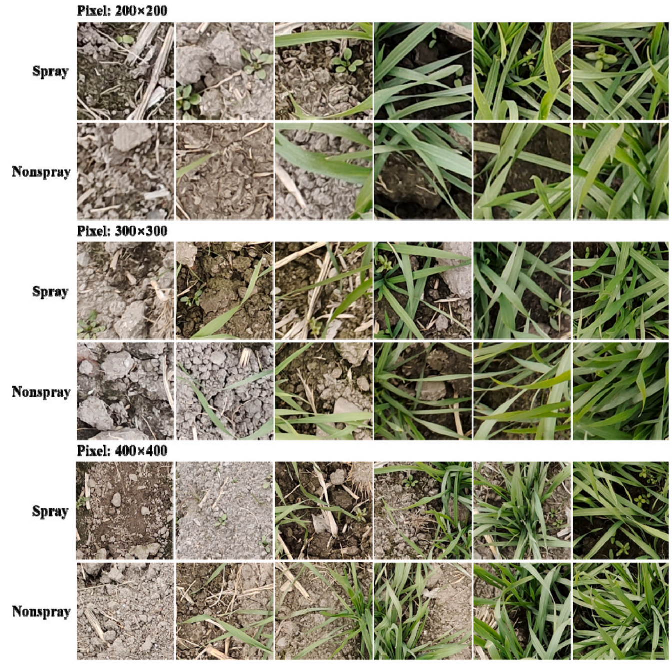

Fig. 1. Examples of images used for training, validation, and testing the weed classification models.

distribution (Shorewala et al., 2021). To the best of our knowledge, the application of SSL in weed detection remains in the nascent stage. The objectives of this research were to (1) investigate the performance of SSL compared to FSL in detecting and locating weeds growing in wheat using training images of varying pixel sizes and labeled data quantities, (2) improve the SSL model to make it more suitable for weed classification tasks, and (3) evaluate the performance of the improved SSL model.

2. Materials and methods

2.1. Dataset preparation

The training and testing images were taken in early December 2020 at two separate wheat fields both measuring 3.5 ha in Yangzhou University Transcultural Science Experiment Station in Yangzhou, Jiangsu, China (32°20′N, 119°23′E). The experimental sites were cultivated with rice (*Oryza sativa* L.) in the previous year. The soil at both experimental locations was classified as a sandy loam, with a pH of 5.8–7.8. The average organic matter content was 16 g kg^{-1}, alkali-hydrolyzed nitrogen was 66.3 mg kg^{-1}, available potassium was 7.2 mg kg^{-1}, and available phosphorus was 59 mg kg^{-1}. After planting wheat for approximately two months, images were randomly captured using a digital camera (Panasonic® DMC-ZS110, Xiamen, China) equipped with a 10X Leica Vario-Elmar Lens (F2.8–5.9 aperture) under automatic exposure settings. The captured images had a resolution of 4300 × 2418 pixels, with a ground sampling resolution of 0.05 cm pixel^{-1}, and were taken between 9:00 a.m. to 5:00 p.m. under various lighting conditions, including clear, cloudy, and partially cloudy skies. Only broadleaf weed species were observed in the two sites including cleavers (*Galium aparine* L.), chickweed (*Malachium aquaticum* L.), Persian speedwell (*Veronica persica* Poir), shepherd's purse [*Capsella bursa-pastoris* (L.) Medik], and sweet woodruff [*Galium odoratum* (L.) Scop.].

The raw images were cropped, resulting in images of 4200 × 2400 and 4000 × 2400 pixels in size. The images of 4200 × 2400 pixels were then divided into 252 sub-images of 200 × 200 pixels and 112 sub-images of 300 × 300 pixels, while the images of 4000 × 2400 pixels were divided into 60 sub-images of 400 × 400 pixels. As shown in Fig. 1, each size of image block was assigned one of two labels: "nonspray" or

Table 1
Training, validation, and testing dataset specifications.

Training				Validation	Testing
Image size	Class	Labeled training	Unlabeled training		
		image quantity			
200 × 200 pixels	Nonspray	100	7000	300	300
	Spray	100		300	300
300 × 300 pixels	Nonspray	100	7000	300	300
	Spray	100		300	300
400 × 400 pixels	Nonspray	100	7000	300	300
	Spray	100		300	300
200 × 200 pixels	Nonspray	200	7000	300	300
	Spray	200		300	300
300 × 300 pixels	Nonspray	200	7000	300	300
	Spray	200		300	300
400 × 400 pixels	Nonspray	200	7000	300	300
	Spray	200		300	300
200 × 200 pixels	Nonspray	300	7000	300	300
	Spray	300		300	300
300 × 300 pixels	Nonspray	300	7000	300	300
	Spray	300		300	300
400 × 400 pixels	Nonspray	300	7000	300	300
	Spray	300		300	300

"spray," with "nonspray" representing sub-images without weeds and thus requiring no herbicide application, while "spray" representing sub-images containing weeds and therefore requiring spot-spaying to achieve precision weed control. To fully evaluate the performance of the proposed method on data with different labels and image sizes, we designed nine datasets, as shown in Table 1. Each dataset includes four categories of data used for labeled training, unlabeled training, validation, and testing. The labeled training and unlabeled training data were utilized to train the model and update the training parameters. The model's hyperparameters were adjusted based on the results of the validation data. Finally, the performance of the model was evaluated using the testing data. The labeled data was divided into 200, 400, and 600 labels for each pixel size. For the same image dataset, the data was augmented with labels derived from the data with fewer labels. This prevented the impact on model training performance due to the inclusion of unrelated label data.

2.2. Proposed method

This section presents a proposed SSL method based on the attention mechanism in detail. The proposed method improved the SSL algorithm Fixmatch (Sohn et al., 2020), combining the methods of consistency regularization and pseudo labels, simplifying the current SSL algorithms and attaining superior performance. To prevent incorrectly labeling unlabeled data with pseudo-labels, we introduced the convolutional block attention module (CBAM) with a hybrid attention mechanism to improve the performance of convolutional neural networks (Woo et al., 2018). CBAM allowed for automatic learning of the most important features for the current task, thereby focusing attention on these features. CBAM comprised a channel attention module and a spatial attention module. When an input feature map $F \in R^{C\times H\times W}$ is an input, the channel attention module performs a 1-dimensional convolution $M_c \in R^{C\times 1\times 1}$, and the output of the convolution is multiplied by the original feature to obtain the output result $F^{'}$. $F^{'}$ is input into spatial attention module, which performs a 2-dimensional convolution $M_c \in$

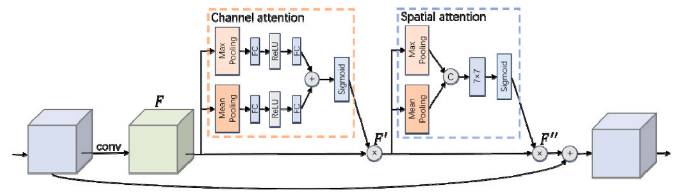

Fig. 2. CBAM integrated with a ResBlock in ResNet50.

$R^{1\times H\times W}$, and the output result is multiplied by the original feature to obtain the feature $F^{''}$. The formulas for calculating $F^{'}$ and $F^{''}$ are as follows:

$$F^{'} = M_c(F) \otimes F$$
$$F^{''} = M_c(F^{'}) \otimes F^{'} \tag{1}$$

Where ⊗ represents element multiplication. Fig. 2 shows the result of adding CBAM to a ResBlock in ResNet50 (He et al., 2016). The figure shows that the input feature map is reduced in dimensionality through two parallel MaxPool layers and AvgPool layers. It then passes through the Share MLP module, where the number of channels is first compressed to 1/r of the original, then expanded to the original number of channels. It then goes through the ReLU activation function to get two activated results. These two output results are element-wisely added and then passed through a sigmoid activation function to obtain the output result of the channel attention module, which is then multiplied by the original feature map. Channel attention module calculates weights for each channel and multiplies them onto the input feature map. In this way, the output feature map retains more important channel information while less important channel information is weakened. The output result of channel attention module then passes through MaxPool and AvgPool layers to obtain two feature maps of size 1 × H × W, which are concatenated through the Concave operation. The resulting feature map is transformed into a single-channel feature map through a 7 × 7 convolution and then passes through a sigmoid to obtain the feature map of spatial attention module, which is finally multiplied by the original feature map. Spatial attention module adjusts the importance of features at different positions by calculating weights for each position and multiplying them onto the input feature map. In this way, more important information in positions is retained in the output, while less important information in positions is weakened.

For binary or multi-class classification, $D_l = \{(x_b, p_b)\}_{b=1}^{B}$ is used as the input for labeled data, where B is the total number of labeled input images, x_b represents a single training image, and p_b is the label corresponding to the image. $D_{ul} = \{u_i\}_{i=1}^{N}$ is used as the input for unlabeled data, where N is the total number of unlabeled input images, and u_i is the unlabeled input image. $H(p, q)$ represents the cross-entropy between the probability distributions p and q, and $p_m(y|x)$ represents the predicted distribution produced by the model for the input x. The proposed methods are continuously used the data augmentation method reported by Sohn et al. (2020). $\mathscr{u}(\cdot)$ and $m(\cdot)$ represent strong and weak data augmentation methods, respectively. For labeled data, normal supervised learning is used for training, and the cross-entropy loss function is used to calculate the labeled loss ℓ_s, according to the following formula:

$$\ell_s = \frac{1}{B}\sum_{b=1}^{B} H(p_b, p_m(y|\mathscr{u}(x_b))) \tag{2}$$

For unlabeled data, $\mathscr{u}(\cdot)$ and $m(\cdot)$ are used to augment the unlabeled samples to predict the augmented samples. The samples enhanced with $\mathscr{u}(\cdot)$ are input into the attention model. When the highest predicted probability of the output result is greater than the given confidence threshold τ, the samples are considered valid and are assigned corresponding pseudo-labels. The samples enhanced with $m(\cdot)$ are input into the attention model, and the cross-entropy loss between the output

Table 2
Values of the hyperparameters for developing the weed classification models.

Deep learning architecture	Optimizer	Learning rate policy	Base learning rate	Weight decay	Training batch size	Evaluation batch size	Momentum
Fullysupervised	AdamW	LambdaLR	5e-5	0.0005	8	16	0.9
Mean teacher	AdamW	LambdaLR	5e-5	0.0005	8	16	0.9
Fixmatch	AdamW	LambdaLR	5e-5	0.0005	8	16	0.9
Proposed method	AdamW	LambdaLR	5e-5	0.0005	8	16	0.9

Abbreviations: Adam, adaptive moment estimation; AdamW, Adam with decoupled weight decay.

result and the samples with pseudo-labels assigned by the corresponding $\omega(\cdot)$ operation is calculated to obtain the unlabeled loss ℓ_u, according to the following formula:

$$\ell_u = \frac{1}{N}\sum_{i=1}^{N} 1(\max\,(p_m(y|\omega(u_i))) > \tau) H(\widehat{p}_m(y|\omega(u_i)), p_m(y|m(u_i))) \tag{3}$$

The pseudo-label of the corresponding u_i sample is denoted as $\widehat{P}_m(y|\omega(u_i))$. By training the model with labeled and unlabeled data, the cross-entropy loss ℓ_s of the labeled data and the cross-entropy loss ℓ_u of the unlabeled data are calculated to obtain the total loss ℓ_t. The total loss is calculated with the following formula, where λ is the weight coefficient of ℓ_u:

$$\ell_t = \ell_s + \lambda_u \ell_u \tag{4}$$

2.3. Experiment design

The proposed method with three baseline methods was examined to verify its effectiveness for weed detection. The first baseline was a FSL method using a standard CNN model with ResNet50 as the backbone network, which was trained solely with labeled data. This method was used to validate the performance of SSL. The second baseline is the mean-teacher model, a SSL method that forms a target-generating teacher model by averaging the model weights based on consistency regularization. This method was used to validate the superiority of SSL models based on consistency regularization and pseudo-labeling. The third baseline was the Fixmatch, which combined consistency regularization and pseudo-labeling to validate the performance of the proposed method. In order to standardize variables and ensure the accuracy of the results, the mean-teacher, Fixmatch, and the proposed methods were all based on the ResNet50 backbone network.

To evaluate the performance of the proposed method, training and validation were conducted on the datasets shown in Table 1. The images had pixels of 200 × 200, 300 × 300, and 400 × 400; for each image size, the number of labeled training images per class was set to 100, 200, and 300, respectively. A total of 7000 images of each pixel size were used as unlabeled training. The test and validation datasets each comprised 300 images per class, totaling 600 images. Of these data, the minimum amount of labeled data (200 labels) accounted for 2.38%, while the maximum amount of labeled data (600 labels) accounted for only 6.82%. The performance of the proposed method on data with the same label but different pixel sizes was verified by comparing the training results of datasets with the same label, but different pixel sizes. Likewise, the performance of the proposed method on data with the same pixel size but different amounts of labels was verified by comparing the training results of datasets with the same pixel size but different amounts of labels.

For both fully-supervised and semi-supervised image classification neural networks, validation, and test results were arranged in a confusion matrix with four possible outcomes, including true positive (tp), false positive (fp), true negative (tn), and false negative (fn). Tp represents the model correctly identifying the target weed; fp represents the model incorrectly predicting the target weed; tn represents the model correctly identifying images without the target weed; fn represents the model failing to predict the real target. The confusion matrix was used to calculate accuracy, precision, recall, and F1 scores. Accuracy measures the overall correctness of the predictions by calculating the ratio of correctly predicted instances (both tp and tn) to the total number of instances. Precision is a metric that measures the proportion of true positive predictions out of all positive predictions made by the system. In the context of sprayer applications, precision indicates how accurately the system identifies the areas that need to be sprayed. A high precision score means the system has a low rate of false positives, i.e., it correctly identifies the areas requiring spraying and minimizes unnecessary spraying in non-target areas. Recall, also known as sensitivity or true positive rate, measures the proportion of true positives predicted by the system out of all actual positive instances in the dataset. In sprayer applications, a recall would indicate how effectively the system detects the areas that require spraying. A high recall score means the system has a low rate of false negatives, i.e., it identifies most areas that need to be sprayed and avoids missing significant portions. The F1 score is a widely used metric for evaluating the performance of classification systems, particularly when dealing with imbalanced datasets. The F1 score provides a balanced view of precision and recall. It helps assess the sprayer system's overall effectiveness by considering both false positives and false negatives. Maximizing the F1 score ensures a good trade-off between precision and recall, aiming for optimal performance in detecting and spraying target areas while minimizing errors. Accuracy, precision, recall, and F1 score were calculated using the following formulas:

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{5}$$

$$Precision = \frac{tp}{tp + fp} \tag{6}$$

$$Recall = \frac{tp}{tp + fn} \tag{7}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

For the validation data, the performance of different models was evaluated using the average precision, average recall, and average F1 score, which were calculated using the following formulas, where C_N was the number of classes, and in this case, $C_N = 2$.

$$Avg\_Pre = \frac{1}{C_N}\sum_{i=1}^{C_N} Precision \tag{9}$$

$$Avg\_Rec = \frac{1}{C_N}\sum_{i=1}^{C_N} Recall \tag{10}$$

$$Avg\_F_1 = \frac{1}{C_N}\sum_{i=1}^{C_N} F_1 \tag{11}$$

The experimental configurations and the hyperparameters used for training the neural networks are presented in Table 2. In order to control the variables between the methods, the training hyperparameters for all methods were unified. The ImageNet database (Deng et al., 2009) was used to pre-train all neural networks. By setting the threshold τ to 0.95, the unlabeled data only gave a pseudo label and joined in the model training if the confidence of the prediction was above the preset threshold. This threshold filtered out images that were incorrectly

Table 3
Image classification neural networks validation and testing results for the detection of broadleaf weed seedlings in wheat.

Deep learning architecture	Labeled	Image size	Validation				Testing			
			Best_Acc	Avg_Pre	Avg_Rec	Avg_F1	Accuracy	Precision	Recall	F1
Fullysupervised	100	200 × 200	0.8200	0.8201	0.8200	0.8200	0.7583	0.7700	0.7370	0.7530
Fullysupervised	100	300 × 300	0.6433	0.6717	0.6433	0.6280	0.6316	0.6500	0.5700	0.6070
Fullysupervised	100	400 × 400	0.7433	0.7458	0.7433	0.7427	0.6100	0.6940	0.3930	0.5020
Meanteacher	100	200 × 200	0.8333	0.8348	0.8333	0.8331	0.7817	0.8240	0.7170	0.7660
Meanteacher	100	300 × 300	0.7033	0.7158	0.7033	0.6990	0.6433	0.5920	0.9200	0.7210
Meanteacher	100	400 × 400	0.8017	0.8058	0.8017	0.8010	0.6580	0.6270	0.7830	0.6960
Fixmatch	100	200 × 200	0.8683	0.8684	0.8683	0.8683	0.8083	0.7810	0.8570	0.8170
Fixmatch	100	300 × 300	0.7267	0.7537	0.7267	0.7192	0.6700	0.6400	0.7770	0.6020
Fixmatch	100	400 × 400	0.8450	0.8529	0.8450	0.8441	0.7250	0.8170	0.5800	0.6780
Proposed method	100	200 × 200	0.9850	0.9850	0.9850	0.9850	0.9600	0.9600	0.9600	0.9600
Proposed method	100	300 × 300	0.8217	0.8512	0.8217	0.8178	0.8160	0.8470	0.7730	0.8080
Proposed method	100	400 × 400	0.9400	0.9401	0.9400	0.9400	0.8417	0.9560	0.7170	0.8190
Fullysupervised	200	200 × 200	0.8667	0.8677	0.8677	0.8666	0.8367	0.8880	0.7700	0.8250
Fullysupervised	200	300 × 300	0.6783	0.7103	0.6783	0.6656	0.6300	0.7070	0.4530	0.5510
Fullysupervised	200	400 × 400	0.8333	0.8351	0.8333	0.8331	0.6400	0.9290	0.3030	0.4570
Meanteacher	200	200 × 200	0.8683	0.8705	0.8683	0.8681	0.8283	0.8660	0.7770	0.8190
Meanteacher	200	300 × 300	0.7500	0.7550	0.7500	0.7488	0.6333	0.6680	0.5300	0.5910
Meanteacher	200	400 × 400	0.8050	0.8068	0.8050	0.8047	0.7200	0.7200	0.7200	0.7200
Fixmatch	200	200 × 200	0.9417	0.9416	0.9416	0.9416	0.8417	0.7920	0.9270	0.8540
Fixmatch	200	300 × 300	0.7550	0.7740	0.7550	0.7507	0.7383	0.7130	0.7970	0.7530
Fixmatch	200	400 × 400	0.9400	0.9424	0.9400	0.9400	0.7550	0.9750	0.5230	0.6810
Proposed method	200	200 × 200	0.9883	0.9883	0.9883	0.9883	0.9633	0.9790	0.9670	0.9630
Proposed method	200	300 × 300	0.8650	0.8880	0.8650	0.8630	0.8250	0.9530	0.6830	0.7960
Proposed method	200	400 × 400	0.9483	0.9489	0.9483	0.9483	0.8717	0.9700	0.7670	0.8570
Fullysupervised	300	200 × 200	0.9417	0.9418	0.9417	0.9417	0.9217	0.9440	0.8970	0.9200
Fullysupervised	300	300 × 300	0.7250	0.7250	0.7250	0.7250	0.6660	0.6540	0.7070	0.6790
Fullysupervised	300	400 × 400	0.9033	0.9064	0.9033	0.9032	0.6550	0.7820	0.4300	0.5550
Meanteacher	300	200 × 200	0.8783	0.8788	0.8783	0.8783	0.8200	0.8900	0.7300	0.8020
Meanteacher	300	300 × 300	0.7533	0.7584	0.7533	0.7521	0.6700	0.6760	0.6530	0.6640
Meanteacher	300	400 × 400	0.8733	0.8754	0.8733	0.8732	0.7733	0.7970	0.7330	0.7640
Fixmatch	300	200 × 200	0.9700	0.9703	0.9700	0.9700	0.9167	0.8770	0.9700	0.9210
Fixmatch	300	300 × 300	0.7650	0.7798	0.7650	0.7619	0.7433	0.7320	0.7670	0.7490
Fixmatch	300	400 × 400	0.9550	0.9561	0.9550	0.9550	0.7680	0.9880	0.5430	0.7010
Proposed method	300	200 × 200	0.9917	0.9917	0.9917	0.9917	0.9800	0.9900	0.9700	0.9800
Proposed method	300	300 × 300	0.9183	0.9241	0.9183	0.9181	0.8583	0.9320	0.7730	0.8450
Proposed method	300	400 × 400	0.9617	0.9619	0.9617	0.9617	0.8767	0.9870	0.7630	0.8610

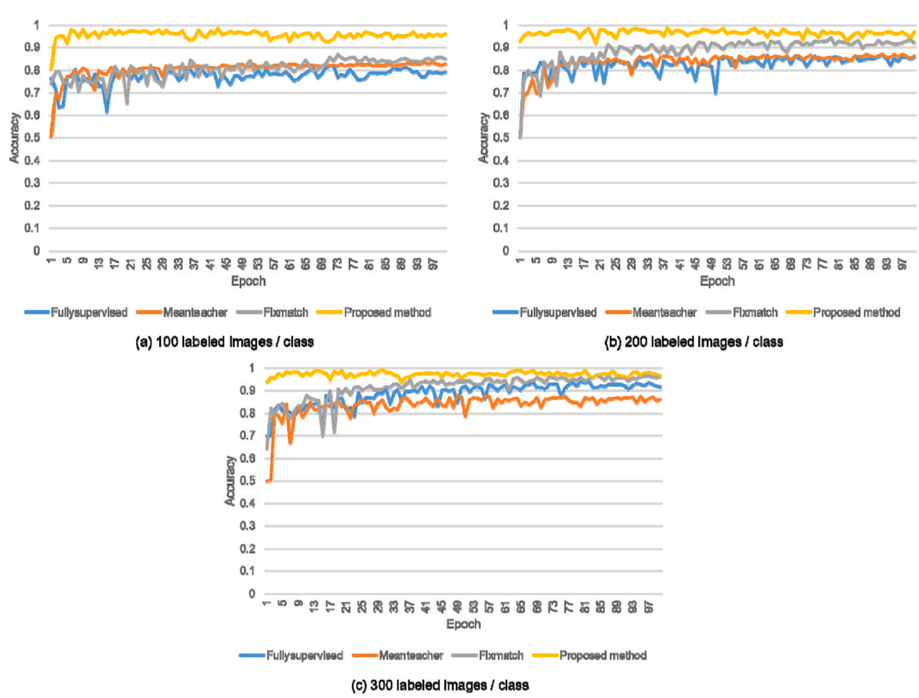

Fig. 3. Comparison of accuracy evaluation of four methods for 200 × 200 pixel images.

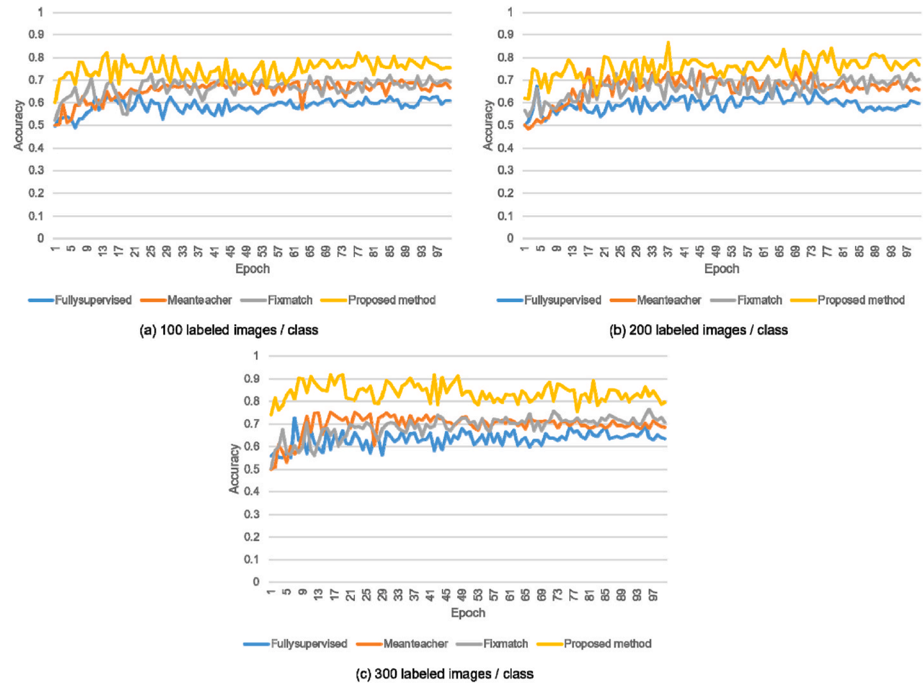

Fig. 4. Comparison of accuracy evaluation of four methods for 300 × 300 pixel images.

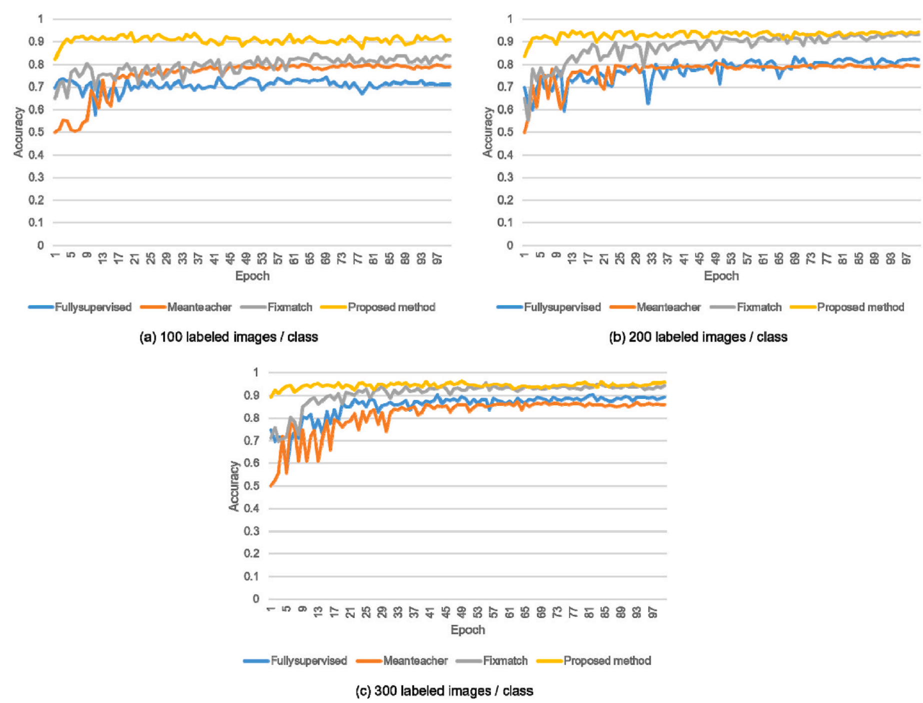

Fig. 5. Comparison of accuracy evaluation of four methods for 400 × 400 pixel images.

assigned pseudo labels, ensuring the accuracy of the model training data. All models were trained on the open-source PyTorch deep learning framework (version 1.8.1, Facebook, San Jose, California, United States) and were tested on an Ubuntu 20.04.1 system with a workstation equipped with an Intel(R) Core(TM) i9-10920X CPU @ 3.50 GHz and an NVIDIA RTX 3080 GPU with 128GB of memory.

3. Results and discussion

In the present study, the SSL and FSL neural networks were trained with different training image sizes and varying numbers of labeled images. The performance of weed detection was evaluated using the same test and validation datasets for training images of the same size, as shown in Table 3. For different sizes of training images, all methods showed excellent performances with an accuracy of ≥0.985 on 200 × 200 pixels labeled training images under conditions of fewer labels. In terms of different numbers of labels, compared to FSL algorithms, the proposed method significantly improved accuracy (≥16.5%) when the neural networks were trained using training images with fewer labels. Overall, the results indicate that the proposed method can reliably

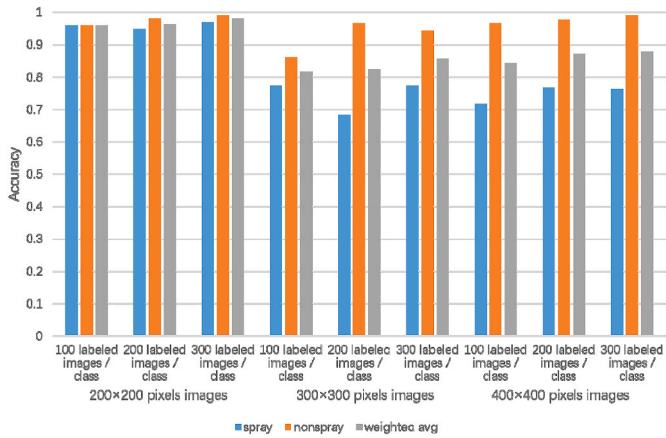

Fig. 6. Accuracy of each class of the proposed method on different datasets.

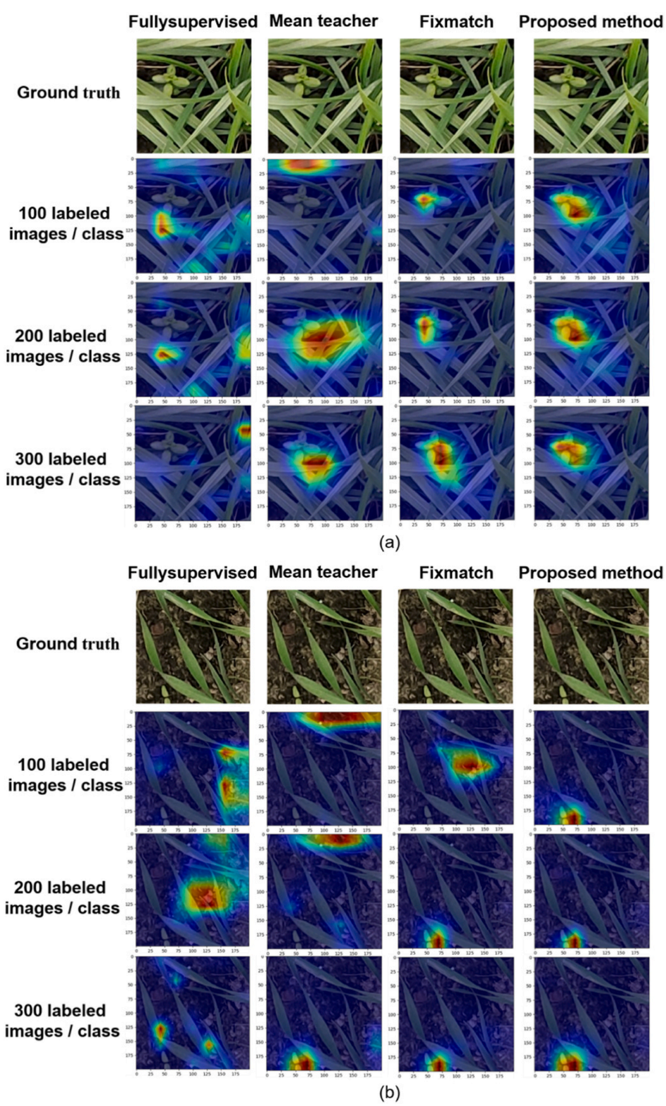

Fig. 7. Grad-CAM interpretability analysis results of four models for 200 × 200 pixel images.

detect broadleaf weed seedlings growing in wheat, compared to FSL algorithms and previous SSL algorithms, with less labeled training data.

The accuracy changes of the four methods on different pixel sizes of the training images are compared in Figs. 3–5. Obviously, after 100 training rounds, the proposed method achieved a higher accuracy compared to the other three methods. After training with 200 × 200 pixel images, the proposed method showed a 16.5%, 12.16%, and 5% increase in accuracy over FSL for 100, 200, and 300 labeled data, respectively. When using 400 × 400 pixel images, the proposed method showed a 19.67%, 11.50%, and 5.84% increase in accuracy, respectively, when compared to FSL developed using 100, 200, and 300 labeled data. The proposed method showed an increase of 17.84%, 18.67%, and 19.33% in accuracy for 300 × 300 pixel images compared to FSL developed using 100, 200, and 300 labeled data. The increase in the number of training images can improve the neural network's performance (Zhuang et al., 2021), resulting in a higher accuracy closer to 1. Therefore, as the number of training images used increases, the improvement in performance of the proposed method compared to the FSL model will be smaller. However, the upward trend observed in the 300 × 300 pixel data suggests that the 300 × 300 pixel image training data is insufficient and requires more data for training. Therefore, the 300 × 300 pixel image is unsuitable for few-labeled SSL tasks.

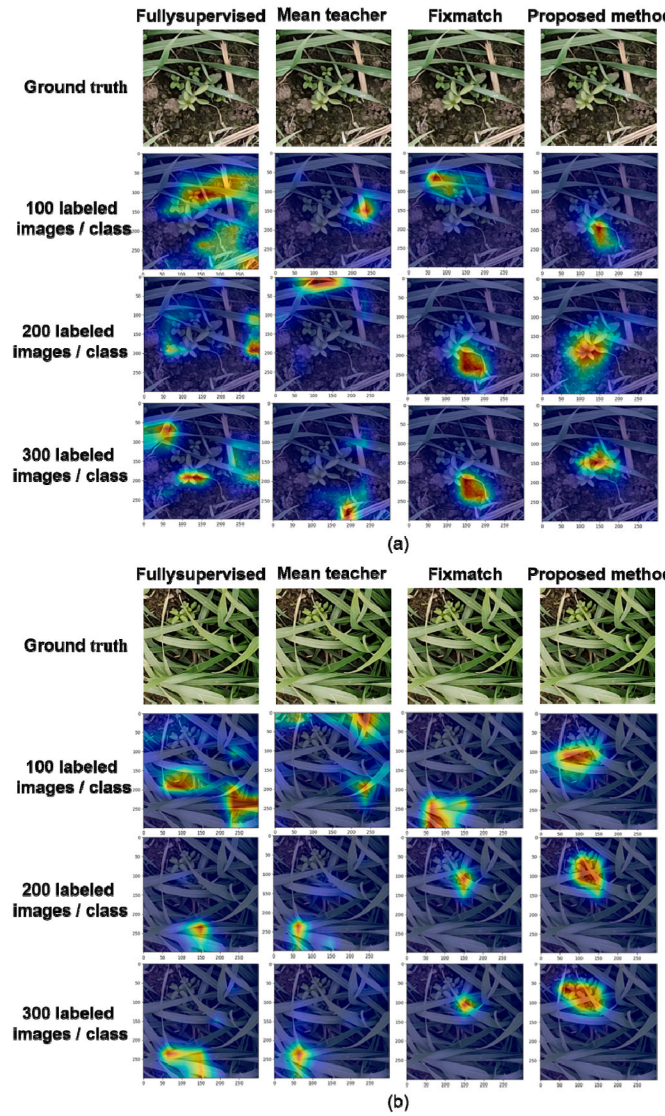

Fig. 8. Grad-CAM interpretability analysis results of four models for 300 × 300 pixel images.

The proposed method and Fixmatch outperform the FSL models in terms of performance on different resolution training images. The Meanteacher method, however, failed to accurately assign correct

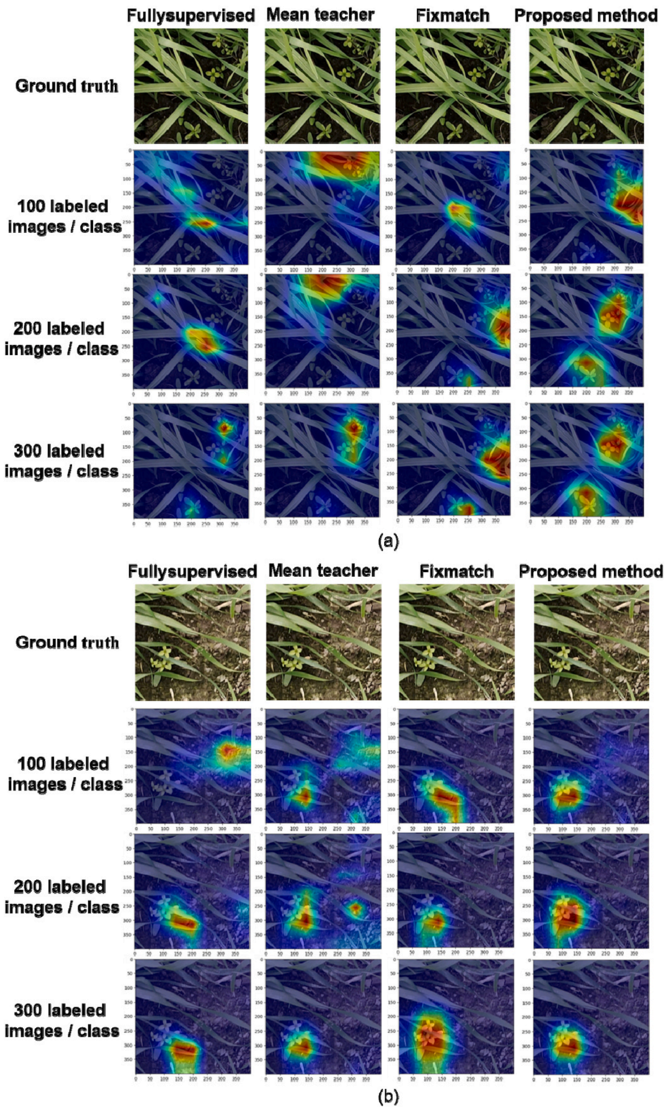

Fig. 9. Grad-CAM interpretability analysis results of four models for 400 × 400 pixel images.

pseudo-labels in the unlabeled data, resulting in a certain amount of noisy samples in the training dataset and thus performing worse than the FSL models on the labeled 600 images with 200 × 200 and 400 × 400 pixels. To obtain a more comprehensive comparison among the four methods, three metrics, including Avg_Pre score, Avg_Rec score, and Avg_F1 score, were compared based on the results at the best epoch of accuracy. The proposed method exhibited the best performance in all metrics for different resolutions and numbers of labeled images. Additionally, due to the limited training data, all models inevitably experienced varying degrees of overfitting on the testing dataset. The proposed method demonstrated the greatest reduction in overfitting when compared to the FSL method and the two SSL methods.

The effectiveness of proposed method is demonstrated in terms of the accuracy for each class in various datasets, as illustrated in Fig. 6. The weighted avg represents the average evaluation metric calculated by weighting according to the number of samples. From the figure, it can be seen that the accuracy for the label "spray" is relatively low (≤77.3%) on 300 × 300 and 400 × 400 pixels images. However, it shows a significant improvement with an accuracy of ≥94.7% on 200 × 200 pixels images. The weighted avg also exhibited a relatively high accuracy (≥96%) compared to 300 × 300 and 400 × 400 pixels images. Furthermore, in order to make the results of the four models more intuitive, two samples from each pixel size of the training data were selected, for a total of six samples, as shown in Figs. 7–9. The attention features and regions of different models were visualized using the gradient-weighted class activation mapping (Grad-CAM) method (Selvaraju et al., 2017). Compared to the FSL and the two SSL methods, the proposed method provided more detailed information and more accurately localized the position of weeds in the image. Therefore, this indicates that the attention mechanism played an important role in the classifier learning process of the proposed method.

To the best of our knowledge, no research has investigated the effect of different pixel images on the performance of the proposed SSL model for weed detection. In previous research, Zhuang et al. (2021) conducted an investigation for FSL models using images with pixel sizes of 200 × 200, 300 × 300, and 400 × 400. The authors evaluated four FSL models, including Densest (Iandola et al., 2014), VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016), and AlexNet (Krizhevsky et al., 2017), achieved F1 scores of ≥0.86; however, this investigation used 22, 000 training images. Despite using only 200 images for training, the proposed method in this study achieved a high F1 score of ≥0.96 for weed classification and detection. Furthermore, the findings of this study suggest 200 × 200 pixel images are more suitable for semi-supervised weed classification than 300 × 300 and 400 × 400 pixel images.

Our results clearly suggest that training SSL models using small-size training images of 200 × 200 pixel offers three advantages: (1) improved weed detection and classification performance compared to models trained with large training image sizes, (2) more precise herbicide spray using smart sprayers, suggesting potential savings in herbicide, and (3) time and labor-efficiencies due to the fact that only six images at a resolution of 4300 × 2418 pixels need to be cropped and manual labeled when preparing the dataset for training the proposed SSL models.

During image processing, the appropriate image size is selected to gridline the images (Jin et al., 2022b). The semi-supervised learning approach is employed to classify the grids, determining the category of each grid. Grids classified as "Spray" indicate the need for herbicide spraying, while those classified as "Nonspray" do not require herbicide spraying. Various postemergence (POST) herbicides, such as sulfonylureas (e.g., floridula, trichoneuron-methyl, and thifensulfuron), synthetic auxins (e.g., 2,4-D, dicamba, MCPA), and photosystem II inhibitors (e.g., metribuzin), are used for POST control of broadleaf weeds in wheat (Baghestani et al., 2008; Curran et al., 2015; Chhokar et al., 2006; Zargar et al., 2019). Smart sprayers can spot-spray these POST herbicides onto the grids containing weeds, thereby reducing herbicide input. The high level of weed detection and classification performance, coupled with the decreased cost of model training, could simplify the development of precision spray systems, leading to a reduction in herbicide inputs without compromising weed control effectiveness in integrated weed management programs.

4. Conclusion

In this research, a SSL algorithm based on an attention mechanism is developed for weed detection in wheat. This is the first effort to introduce a SSL algorithm for weed detection in wheat, which has been improved based on the existing semi-supervised model. This has greatly reduced the impact of noisy samples, thus giving the proposed method a stronger ability to capture weed features. The application of the SSL algorithm developed in this study could significantly contribute to the detection of weeds growing in wheat or other small grain crops. The proposed method, which is based on average consistency regularization and pseudo labels, trained the model using both labeled and unlabeled data. The introduction of the mixed attention mechanism CBAM automatically learns the relative importance of channels and positions, thus improving the model's performance. Furthermore, the proposed method is significantly more effective than an FSL model. Training the model with 200 × 200 pixel images is more suitable for the semi-supervised weed classification task than using 300 × 300 or 400 × 400 pixel images. Compared to previous SSL methods, the proposed SSL method can greatly reduce the time and manual cost of data labeling, thus simplifying the development of weed classifiers. To further improve the performance of weed detection, an additional study will explore extended SSL model algorithms, such as CRMatch (Fan et al., 2023). In the present study, the captured images in the experimental sites included only broadleaf weed species. Further studies are needed to evaluate the performances of these techniques in detecting and discriminating newly emerged grass weeds growing in wheat.

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.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 32072498).

References

Ahmad, A., Saraswat, D., Aggarwal, V., Etienne, A., Hancock, B., 2021. Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. Comput. Electron. Agric. 184, 106081 https://doi.org/10.1016/j.compag.2021.106081.

Ahmad Loti, N.N., Mohd Noor, M.R., Chang, S.W., 2021. Integrated analysis of machine learning and deep learning in chili pest and disease identification. J. Sci. Food Agric. 101 (9), 3582–3594. https://doi.org/10.1002/jsfa.10987.

Andrea, C.C., Daniel, B.B.M., Misael, J.B.J., 2017. Precise weed and maize classification through convolutional neuronal networks. In: 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), pp. 1–6. https://doi.org/10.1109/ETCM.2017.8247469. Salinas, Ecuador.

Asseng, S., Foster, I.A.N., Turner, N.C., 2011. The impact of temperature variability on wheat yields. Global Change Biol. 17 (2), 997–1012. https://doi.org/10.1111/j.1365-2486.2010.02262.x.

Baghestani, M.A., Zand, E., Soufizadeh, S., Beheshtian, M., Haghighi, A., Barjasteh, A., Birgani, G.D., Deihimfard, R., 2008. Study on the efficacy of weed control in wheat (*Triticum aestivum L.*) with tank mixtures of grass herbicides with broadleaved herbicides. Crop Protect. 27, 104–111. https://doi.org/10.1016/j.cropro.2007.04.013.

Bakhshipour, A., Jafari, A., Nassiri, S.M., Zare, D., 2017. Weed segmentation using texture features extracted from wavelet sub-images. Biosyst. Eng. 157, 1–12. https://doi.org/10.1016/j.biosystemseng.2017.02.002.

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. https://doi.org/10.1016/j.compag.2017.12.032.

Bàrberi, P., 2002. Weed management in organic agriculture: are we addressing the right issues? Weed Res. 42 (3), 177–193. https://doi.org/10.1046/j.1365-3180.2002.00277.x.

Bernotas, G., Scorza, L.C.T., Hansen, M.F., Hales, I.J., Halliday, K.J., Smith, L.N., Smith, M.L., McCormick, A.J., 2019. A photometric stereo-based 3D imaging system using computer vision and deep learning for tracking plant growth. GigaScience 8 (5), giz056. https://doi.org/10.1093/gigascience/giz056.

Berthelot, D., Carlini, N., Cubuk, E.D., Kurakin, A., Sohn, K., Zhang, H., Raffel, C., 2019a. Rematch: semi-supervised learning with distribution alignment and augmentation anchorings. arXiv preprint arXiv:1911.09785. https://doi.org/10.48550/arXiv.1911.09785.

Berthelot, D., Carlini, N., Goodfellow, J.I., Papernot, N., Oliver, A., Raffelet, C., 2019b. Dispatch: a holistic approach to semi-supervised learning. Adv. Neural Inf. Process. Syst. 32 https://doi.org/10.48550/arXiv.1905.02249.

Burks, T.F., Shearer, S.A., Payne, F.A., 2000. Classification of weed species using color texture features and discriminant analysis. Trans. ASAE (Am. Soc. Agric. Eng.) 43 (2), 441–448. https://doi.org/10.13031/2013.2723.

Chandy, A., 2019. Pest infestation identification in coconut trees using deep learning. Int. J. Artif. Intell. 1 (1), 10–18. https://doi.org/10.36548/jaicn.2019.1.002.

Chen, S., Bortsova, G., García-Uceda Juárez, A., Van Tulder, G., De Bruijne, M., 2019. Multi-task attention-based semi-supervised learning for medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, pp. 457–465. https://doi.org/10.1007/978-3-030-32248-9_51.

Cheng, X., Zhang, Y., Chen, Y., Wu, Y., Yue, Y., 2017. Pest identification via deep residual learning in complex background. Comput. Electron. Agric. 141, 351–356. https://doi.org/10.1016/j.compag.2017.08.005.

Chhokar, R., Sharma, R., Chauhan, D., Mongia, A., 2006. Evaluation of herbicides against Phalaris minor in wheat in north-western Indian plains. Weed Res. 46 (1), 40–49. https://doi.org/10.1111/j.1365-3180.2006.00485.x.

Curran, W.S., Wallace, J.M., Mirsky, S., Crockett, B., 2015. Effectiveness of herbicides for control of hairy vetch (*Vicia villosa*) in winter wheat. Weed Technol. 29 (3), 509–518. https://doi.org/10.1614/WT-D-14-00139.1.

Curtis, T., Halford, N.G., 2014. Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. Ann. Appl. Biol. 164 (3), 354–372. https://doi.org/10.1111/aab.12108.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., F, F., Li, 2009. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. https://doi.org/10.1109/CVPR.2009.5206848. Miami, FL, USA, 2009.

Dos Santos Ferreira, A., Matte Freitas, D., Goncalves da Silva, G., Pistori, H., Theophilo Folhes, M., 2017. Weed detection in soybean crops using Convects. Comput. Electron. Agric. 143, 314–324. https://doi.org/10.1016/j.compag.2017.10.027.

Fan, Y., Kukleva, A., Dai, D., Schiele, B., 2023. Revisiting consistency regularization for semi-supervised learning. Int. J. Comput. Vis. 131, 626–643. https://doi.org/10.1007/s11263-022-01723-4.

Franz, E., Gebhardt, M.R., Unklesbay, K.B., 1991. The use of local spectral properties of leaves as an aid for identifying weed seedlings in digital images. Trans. ASAE (Am. Soc. Agric. Eng.) 34 (2), 682–687. https://doi.org/10.13031/2013.31717.

Gao, Q., Huang, Y., Gao, X., Shen, W., Zhang, H., 2015. A novel semi-supervised learning for face recognition. Neurocomputing 152, 69–76. https://doi.org/10.1016/j.neucom.2014.11.018.

Guo, J., Wang, Q., Li, Y., 2021. Semi-supervised learning based on convolutional neural network and uncertainty filter for façade defects classification. COMPUT-AIDED CIV INF. 36 (3), 302–317. https://doi.org/10.1111/mice.12632.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Heap, I., 2023. The international herbicide-resistant weed database, 2023. www.weedscience.org. (Accessed 20 January 2023).

Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., Keutzer, K., 2014. Densenet: implementing efficient convenes descriptor pyramids. arXiv preprint arXiv:1404.1869. https://doi.org/10.48550/arXiv.1404.1869.

Igrejas, G., Branlard, G., 2020. The Importance of Wheat, Wheat Quality for Improving Processing and Human Health. Springer. Cham, pp. 1–7. https://doi.org/10.1007/978-3-030-34163-3_1.

Jin, X., Bagavathiannan, M., McCullough, P.E., Chen, Y., Yu, J., 2022b. A deep learning based method for classification, detection, and localization of weeds in turfgrass. Pest Manag. Sci. 78 (11), 4809–4821. https://doi.org/10.1002/ps.7102.

Jin, X., Che, J., Chen, Y., 2021. Weed identification using deep learning and image processing in vegetable plantation. IEEE Access 9, 10940–10950. https://doi.org/10.1109/ACCESS.2021.3050296.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Kazmi, W., Garcia-Ruiz, F., Nielsen, J., Rasmussen, J., Andersen, H.J., 2015. Exploiting affine invariant regions and leaf edge shapes for weed detection. Comput. Electron. Agric. 118, 290–299. https://doi.org/10.1016/j.compag.2015.08.023.

Kennedy, H.J., Fennimore, S.A., Slaughter, D.C., Nguyen, T.T., Vuong, V.L., Raja, R., Smith, R.F., 2020. Crop signal markers facilitate crop detection and weed removal from lettuce and tomato by an intelligent cultivator. Weed Technol. 34 (3), 342–350. https://doi.org/10.1017/wet.2019.120.

Khaki, S., Wang, L., 2019. Crop yield prediction using deep neural networks. Front. Plant Sci. 10, 621. https://doi.org/10.3389/fpls.2019.00621.

Khan, I., Jan, A.U., Gul, F., Ali, K., Ghaffar, A., Ahmed, S., Ahmad, N., 2011. Weeds as a major constraint in wheat production in district Peshawar. Pak. J. Weed Sci. Res. 17 (4).

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.

Laine, S., Aila, T., 2016. Temporal assembling for semi-supervised learning. arXiv preprint arXiv:1610.02242. https://doi.org/10.48550/arXiv.1610.02242.

Lin, F., Zhang, D., Huang, Y., Wang, X., Chen, X., 2017. Detection of corn and weed species by the combination of spectral, shape and textural features. Sustainability 9 (8), 1335. https://doi.org/10.3390/su9081335.

Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. Comput. Electron. Agric. 63 (2), 282–293. https://doi.org/10.1016/j.compag.2008.03.009.

Miyato, T., Dai, A.M., Goodfellow, I., 2016. Adversarial training methods for semi-supervised text classification. arXiv preprint arXiv:1605.07725. https://doi.org/10.48550/arXiv.1605.07725.

Murmu, S., Biswas, S., 2015. Application of fuzzy logic and neural network in crop classification: a review. Aquat. Procedia. 4, 1203–1210. https://doi.org/10.1016/j.aqpro.2015.02.153.

Nevavuori, P., Narra, N., Lipping, T., 2019. Crop yield prediction with deep convolutional neural networks. Comput. Electron. Agric. 163, 104859 https://doi.org/10.1016/j.compag.2019.104859.

Nong, C., Fan, X., Wang, J., 2022. Semi-supervised learning for weed and crop segmentation using UAV imagery. Front. Plant Sci. 13 https://doi.org/10.3389/fpls.2022.927368.

Norsworthy, J.K., Ward, S.M., Shaw, D.R., Llewellyn, R.S., Nichols, R.L., Webster, T.M., Bradley, K.W., Frisvold, G., Powles, S.B., Burgos, N.R., Witt, W.W., Barrett, M., 2012. Reducing the risks of herbicide resistance: best management practices and recommendations. Weed Sci. 60 (SP1), 31–62. https://doi.org/10.1614/ws-d-11-00155.1.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D., Walsh, J., 2020. Deep learning vs. traditional computer vision. Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC) 11. https://doi.org/10.1007/978-3-030-17795-9_10. SpringerPlus. 128–144.

Pérez-Ortiz, M., Peña, J.M., Gutiérrez, P.A., Torres-Sánchez, J., Hervás-Martínez, C., López-Granados, F., 2015. A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. Appl. Soft Comput. 37, 533–544. https://doi.org/10.1016/j.asoc.2015.08.027.

Peterson, M.A., Collavo, A., Ovejero, R., Shivrain, V., Walsh, M.J., 2018. The challenge of herbicide resistance around the world: a current summary. Pest Manag. Sci. 74 (10), 2246–2259. https://doi.org/10.1002/ps.4821.

Razfar, N., True, J., Bassiouny, R., Venkatesh, V., Kashef, R., 2022. Weed detection in soybean crops using custom lightweight deep learning models. J. Agric. Food Res. 8, 100308 https://doi.org/10.1016/j.jafr.2022.100308.

Rizve, M.N., Duarte, K., Rawat, Y.S., Shah, M., 2021. In defense of pseudo-labeling: an uncertainty-aware pseudo-label selection framework for semi-supervised learning. arXiv preprint arXiv:2101.06329. https://doi.org/10.48550/arXiv.2101.06329.

Scudder, H., 1965. Probability of error of some adaptive pattern-recognition machines. IEEE Trans. Inf. Theor. 11 (3), 363–371. https://doi.org/10.1109/TIT.1965.1053799.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-cam: visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision. https://doi.org/10.48550/arXiv.1610.02391, 618–626.

Shirzadifar, A., Bajwa, S., Nowatzki, J., Shojaeiarani, J., 2020. Development of spectral indices for identifying glyphosate-resistant weeds. Comput. Electron. Agric. 170, 105276 https://doi.org/10.1016/j.compag.2020.105276.

Shorewala, S., Ashfaque, A., Sidharth, R., Verma, U., 2021. Weed density and distribution estimation for precision agriculture using semi-supervised learning. IEEE Access 9, 27971–27986. https://doi.org/10.1109/ACCESS.2021.3057912.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-Scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556.

Søgaard, A., 2013. Semi-supervised learning and domain adaptation in natural language processing. Synth. Met. 6 (2), 1–103. https://doi.org/10.1007/978-3-031-02149-7.

Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A., Li, C.-L., 2020. Fixmatch: simplifying semi-supervised learning with consistency and confidence. Adv. Neural Inf. Process. Syst. 33, 596–608. https://doi.org/10.48550/arXiv.2001.07685.

Tang, L., Tian, L., Steward, B.L., 2000. Color image segmentation with genetic algorithm for in-field weed sensing. Trans. ASAE (Am. Soc. Agric. Eng.) 43 (4), 1019–1027. https://doi.org/10.13031/2013.2970.

Torres-Sánchez, J., López-Granados, F., Pena, J.M., 2015. An automatic object-based method for optimal thresholding in UAV images: application for vegetation detection in herbaceous crops. Comput. Electron. Agric. 114, 43–52. https://doi.org/10.1016/j.compag.2015.03.019.

Van Engelen, J.E., Hoos, H.H., 2020. A survey on semi-supervised learning. Mach. Learn. 109 (2), 373–440. https://doi.org/10.1007/s10994-019-05855-6.

Van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: a systematic literature review. Comput. Electron. Agric. 177, 105709 https://doi.org/10.1016/j.compag.2020.105709.

Wang, D., Zhang, Y., Zhang, K., Wang, L., 2020. Focalis: semi-supervised learning for 3d medical image detection. proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition 3951–3960.

Wang, H., Liu, W., Jin, T., Peng, X., Zhang, L., Wang, J., 2020. Bipyrazone: a new HPPD-inhibiting herbicide in wheat. Sci. Rep. 10, 5521. https://doi.org/10.1038/s41598-020-62116-6.

Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A., 1995. Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASAE (Am. Soc. Agric. Eng.) 38 (1), 259–269. https://doi.org/10.13031/2013.27838.

Woo, S., Park, J., Lee, J.Y., 2018. Cam: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 3–19. https://doi.org/10.48550/arXiv.1807.06521.

Wu, L., Wen, Y., 2009. Weed/corn seedling recognition by support vector machine using texture features. Afr. J. Agric. Res. 4 (9), 840–846.

Wu, Z., Chen, Y., Zhao, B., Kang, X., Ding, Y., 2021. Review of weed detection methods based on computer vision. SNSR 21 (11), 3647. https://doi.org/10.3390/s21113647.

Xiao, J., Chen, L., Pan, F., Deng, Y., Ding, C., Liao, M., Su, X., Cao, H., 2020. Application method affects pesticide efficiency and effectiveness in wheat fields. Pest Manag. Sci. 76 (4), 1256–1264. https://doi.org/10.1002/ps.5635.

Yan, J., Wang, X., 2022. Unsupervised and semi-supervised learning: the next frontier in machine learning for plant systems biology. Plant J. 111 (6), 1527–1538. https://doi.org/10.1111/tpj.15905.

Yasrab, R., Zhang, J., Smyth, P., Pound, M.P., 2021. Predicting plant growth from time-series data using deep learning. Rem. Sens. 13 (3), 331. https://doi.org/10.3390/rs13030331.

Yin, C., Xiang, J., Zhang, H., Wang, J., Yin, Z., Kim, J.-U., 2015. A new SVM method for short text classification based on semi-supervised learning. In: 2015 4th International Conference on Advanced Information Technology and Sensor Application (AITS), pp. 100–103. https://doi.org/10.1109/AITS.2015.34. Harbin, China.

Yu, J., Schumann, A.W., Cao, Z., Sharpe, S.M., Boyd, N.S., 2019a. Weed detection in perennial ryegrass with deep learning convolutional neural network. Front. Plant Sci. 10, 1422. https://doi.org/10.3389/fpls.2019.01422.

Yu, J., Sharpe, S.M., Schumann, A.W., Boyd, N.S., 2019b. Deep learning for image-based weed detection in turfgrass. Eur. J. Agron. 104, 78–84. https://doi.org/10.1016/j.eja.2019.01.004.

Yu, J., Sharpe, S.M., Schumann, A.W., Boyd, N.S., 2019c. Detection of broadleaf weeds growing in turfgrass with convolutional neural networks. Pest Manag. Sci. 75 (8), 2211–2218. https://doi.org/10.1002/ps.5349.

Zargar, M., Bayat, M., Lyashko, M., Chauhan, B., 2019. Postemergence herbicide applications impact Canada thistle control and spring wheat yields. Agron. J. 111 (6), 2874–2880. https://doi.org/10.2134/agronj2019.02.0125.

Zhang, B., Wang, Y., Hou, W., Wu, H., Wang, J., Okumura, M., Shinozaki, T., 2021. Rematch: boosting semi-supervised learning with curriculum pseudo labeling. Adv. Neural Inf. Process. Syst. 34, 18408–18419.

Zhao, C., He, D., Qiao, Y., 2013. Identification method of multi-feature weed based on multi-spectral images and data mining. TCSAE 29 (2), 192–198. https://doi.org/10.3969/j.issn.1002-6819.2013.02.027.

Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. Remote Sens. Environ. 221, 430–443. https://doi.org/10.1016/j.rse.2018.11.032.

Zhuang, J., Li, X., Bagavathiannan, M., Jin, X., Yang, J., Meng, W., Li, T., Li, L., Wang, Y., Chen, Y., 2021. Evaluation of different deep convolutional neural networks for detection of broadleaf weed seedlings in wheat. Pest Manag. Sci. 78 (2), 521–529. https://doi.org/10.1002/ps.6656.