Research Article

Received: 19 July 2021 Revised: 19 September 2021 Accepted article published: 24 September 2021 Published online in Wiley Online Library: 5 October 2021

(wileyonlinelibrary.com) DOI 10.1002/ps.6656

Evaluation of different deep convolutional neural networks for detection of broadleaf weed seedlings in wheat

Jiayao Zhuang,[a] Xuehan Li,[a] Muthukumar Bagavathiannan,[b] Xiaojun Jin,[c] Jie Yang,[c] Wenting Meng,[a] Tao Li,[a] Lanxi Li,[a] Yundi Wang,[d] Yong Chen[c] and Jialin Yu[a,b*]

Abstract

BACKGROUND: In-field weed detection in wheat (*Triticum aestivum* L.) is challenging due to the occurrence of weeds in close proximity with the crop. The objective of this research was to evaluate the feasibility of using deep convolutional neural networks for detecting broadleaf weed seedlings growing in wheat.

RESULTS: The object detection neural networks, including CenterNet, Faster R-CNN, TridenNet, VFNet, and You Only Look Once Version 3 (YOLOv3) were insufficient for weed detection in wheat because the recall never exceeded 0.58 in the testing dataset. The image classification neural networks including AlexNet, DenseNet, ResNet, and VGGNet were trained with small (5500 negative and 5500 positive images) or large training datasets (11 000 negative and 11 000 positive images) and three training image sizes (200 × 200, 300 × 300, and 400 × 400 pixels). For the small training dataset, increasing image sizes decreased the F1 scores of AlexNet and VGGNet but generally increased the F1 scores of DenseNet and ResNet. For the large training dataset, no obvious difference was detected between the training image sizes since all neural networks exhibited remarkable classification accuracies with high F1 scores (≥0.96). All image classification neural networks exhibited high F1 scores (≥0.99) when trained with the large training dataset and the training images of 200 × 200 pixels.

CONCLUSION: CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3 were insufficient, while AlexNet, DenseNet, ResNet, and VGGNet trained with a large training dataset were highly effective for detection of broadleaf weed seedlings in wheat.

Keywords: artificial intelligence; deep learning; machine learning; digital agriculture; site-specific weed management

1 INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important crops in the world and contributes daily sources of dietary calories for a large proportion of the world's population.[1] Weed competition is a serious constraint for wheat production worldwide.[2] For example, a density of 10 intra-row white mustard (*Sinapis alba* L.) plants per square meter reduced wheat yield up to 32%.[3] Spraying herbicides is the most commonly employed strategy for weed control,[4,5] although the cases of herbicide resistance in wheat have been increasingly reported over the past few decades.[6] Weeds typically show a random, patchy distribution in fields, but postemergence (POST) herbicides are broadcast-applied for weed control uniformly throughout the field, leading to herbicide applications in areas where it is not necessary.

Precision herbicide application, based on an accurate, reliable, and automatic weed detection technology, can substantially reduce herbicide input and weed control costs.[7,8] Previous researchers explored a variety of sensing methods, such as fluorescence,[9–11] visible or near-infrared spectroscopy,[12,13] hyper- or multi-spectral imaging,[14,15] and machine vision,[16–18] for weed detection. Nevertheless, the introduction of smart sprayers into practical farming, particularly for wheat, is still lacking. The major obstacle limiting the widespread adoption of automatic weed control is the absence of robust sensing technology to provide reliable weed detection.[7] In wheat, weed detection is especially difficult due to the presence of a variety of weed species growing in close proximity with the crop.

In recent years, machine learning, particularly deep convolutional neural networks (DCNNs) that are used in conjunction with

* Correspondence to: J Yu, Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843, USA. E-mail: yu.jialin@tamu.edu

a Co-Innovation Center for sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

b Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA

c College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing, China

d Department of Computer Science, Stevens Institute of Technology, Hoboken, NJ, USA

a graphics processing unit (GPU), has demonstrated remarkable capability in various scientific applications, such as self-driving cars,[19] speed recognition,[20] natural language processing,[21] and precision medicine.[22] DCNNs exhibit a tremendous ability to learn and extract complex features from images[19] and have recently emerged as a powerful tool in various agricultural domains, such as crop yield prediction,[23] counting fruit number,[24] and plant phenotyping.[25] For example, Ghosal *et al.*[25] developed DCNNs model that can reliably classify and identify abiotic (chemical injury and nutrient deficiency) and biotic (bacterial and fungal diseases) stresses in soybean [*Glycine max* (L.) Merr.].

The potential for DCNNs for real-time and reliable detection of weeds has demonstrated in various cropping systems such as corn (*Zea mays* L.),[26] dormant or actively growing turfgrass,[17,18] potato (*Solanum tuberosum* L.),[27] plastic-mulched vegetables,[28] sunflower (*Helianthus annuus* L.),[26] and soybean.[29] Although the endeavor of using DCNNs for weed detection is relatively new, previous studies achieved high accuracies (≥ 98%) for classifying images containing different turfgrass species or weeds growing in turfgrass,[18,30] and for detecting broadleaf and grass weeds growing in soybean.[29] Recently, Su *et al.*[31] proposed a DCNNs-based method that can provide real-time detection of inter-row ryegrass (*Lolium* spp.) in wheat. Unfortunately, the detection of intra-row weeds, especially at the seedling growth stage, is still a challenging task in wheat.

The capability of machine learning models for detecting weeds at the seedling growth stage is important because POST herbicides are typically more effective for controlling weeds at the earlier rather than later stages.[32] In addition, herbicide rates can be adjusted according to weed growth stages to achieve better control.[33] However, the feasibility of using DCNNs for detection of weeds at the seedling growth stage of wheat has not been well explored. In this study, we investigate the potential for using DCNNs-based methods for detection of broadleaf weed seedlings in wheat.

2 MATERIALS AND METHODS

2.1 Image acquisition

Training and testing images were taken multiple times at two separate fields in Yangzhou University Pratacultural Science Experiment Station, Yangzhou, Jiang Su, China (32°20′N, 119°23′E) from December 3, 2020 to December 12, 2020. For both fields, where the images were acquired, cleavers (*Galium aparine* L.), crickweed (*Malachium aquaticum* L.), and shepherd's purse [*Capsella bursa-pastoris* (L.) Medik] were the major broadleaf weed species observed. Images were collected approximately 2 months after planting wheat. Images (4300 × 2418 pixels) were acquired using a digital camera (Panasonic® DMC-ZS110 Xiamen, Fujian, China) equipped with 10X Leica Vario-Elmarit Lens (F2.8–5.9 aperture) at an auto-exposure setting. Images were acquired at a ground sampling resolution of 0.05 cm pixel^{-1} and were acquired during daytime from 9:00 a.m. to 5:00 p.m. under various lighting conditions such as clear, cloudy, and partially cloudy skies.

2.2 Object detection

For training object detection neural networks, all training and testing images were cropped to 2560 × 1440 pixels using Irfanview (version 5.50, Irfan Skijan, Jaice, Bosnia). A total of 906 images containing broadleaf weeds growing in wheat were used in the training dataset, 90 (10%) of which were randomly selected as the validation dataset. A total of 30 images containing weeds while growing in wheat were randomly selected and used as the testing dataset (Table 1).

The object detection architectures investigated in the present study included CenterNet,[34] Faster R-CNN,[35] Trident Network (TridentNet),[36] Variable Filter Net (VFNet),[37] and You Only Look Once (YOLO) version 3 (YOLOv3),[38] CenterNet utilizes keypoint estimation to identify center points and meanwhile regresses to all other object properties, including three-dimensional (3D) location, even pose, orientation, and size.[34] Faster R-CNN shares the convolutional features of Region Proposal Network and Fast R-CNN and thus enables a near real-time frame detection speed.[35] TridentNet can generate scale-specific feature maps using a uniform representational power.[36] TridentNet was constructed using a parallel multi-branch architecture in which each branch has different receptive fields but shares the same transformation parameters.[36] VFNet captures a hierarchy of features through the application of variable filter sizes along with the audio spectrograms.[37] VFNet was originally designed for accent classification[37] but its feasibility for weed detection was evaluated in the present study. YOLOv3 was developed based on YOLO[39] and YOLO version 2.[40] YOLO is a single-stage deep learning architecture that utilizes independent logistic classifiers and generates multi-labeled bounding boxes.[38] The inference time was the priority when YOLO was designed.[38]

The image area containing weeds was annotated using LabelImg (an open source software available at https://github.com/tzutalin/labelImg). The object detection neural networks detect the targets based on bounding box coverage. The numbers of bounding box annotations were 5261, 510, and 373 for training, validation, and testing datasets, respectively. The bounding box annotation was chosen rather than pixel-wise annotation in this study due to reduced time requirement and increased detection accuracy, as indicated by Sa *et al.*[41]

CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3 were pre-trained with the COCO dataset.[42] YOLOv3 was trained and tested with the Darknet (an open source neural network

Table 1. Hyper-parameters used for training object detection neural networks

Parameters	CenterNet2	Faster R-CNN	TridentNet	VFNet	YOLOv3
Batch	1	4	2	4	64
Momentum	0.90	0.90	0.90	0.90	0.90
Decay	0.0001	0.0001	0.0001	0.0001	0.0005
Learning rate	0.04	0.02	0.02	0.01	0.001
Policy	Step	Step	Step	Step	Steps
Epoch	120	24	24	24	30

framework available at https://pjreddie.com/darknet/),[39] while the training and testing of CenterNet, Faster R-CNN, TridentNet, and VFNet were performed using mmDetection based on the Pytorch framework (an open source deep learning framework available at https://pytorch.org/; Facebook, San Jose, CA, USA). For all models, training and testing were performed on a GeForce RTX 2080 Ti with 64 GB of memory. During training, an intersection over union (IoU) measuring the overlap ratio between the predicted and actual bounding boxes was employed to estimate if the object detected was a true positive, with a threshold of 0.5. Training continued until the average loss error ceased to decrease further or the desirable parameters such as mean average precision (mAP), precision, or recall did not increase any further.

2.3 Image classification

The image classification neural networks including AlexNet,[43] DenseNet,[44] Residual Network (ResNet),[45] and VGGNet[46] were selected for evaluating the feasibility of using DCNNs for detection of broadleaf weed seedlings growing in wheat. AlexNet consisted of eight layers, including three fully connected layers and five convolutional layers.[43] DenseNet computes multi-scale features from the convolutional layers of DCNN-based object classifier.[44] DenseNet computes descriptors densely without any regard for region of proposal windows.[44] ResNet consisted of 18-layer and 34-layer residual nets.[45] Instead of learning unreferenced functions, ResNet was constructed by reformulating the layers as learning residual functions with reference to the layer inputs.[45] VGGNet used in the present study is composed of 16 weight layers.[46] VGGNet utilizes a stack of convolutional layers followed by max-poling layers.[46]

All training and testing images were cropped into sub-images with a resolution of 200×200, 300×300, or 400×400 pixels using Irfanview. The quantity of images for the training, validation, and testing datasets was standardized in order to compare the results of all neural networks. A total of 24 image classification neural networks were trained and tested. The neural networks were trained with two separate training datasets containing different amounts of training images. The small training dataset contained a total of 5500 negative (with no weeds) and 5500 positive images (with weeds), while the large training dataset contained a total of 11 000 negative and 11 000 positive images. The validation or testing dataset consisted of 150 positive and 150 negative sub-images.

Training and testing were performed on a GeForce RTX 2080 Ti with 64 GB of memory using the PyTorch open source deep learning framework. All neural networks were pre-trained using the ImageNet database.[47] To achieve a fair composition between the neural networks, all neural networks were converted to the PyTorch version by modifying the weights to the corresponding PyTorch compositions. The hyper-parameters employed across the experimental configurations for training are presented in Table 2.

2.4 Evaluation

For both object detection and image classification neural networks, the validation and testing results are arranged in a confusion matrix with four possible outcomes, including true positive (tp), false positive (fp), true negative (tn), and false negative (fn). In this context, a tp represents the network correctly identified the target weeds; a fp represents the network incorrectly predicted the target weeds; a tn represents the network correctly identified the images without the target weeds; and a fn represents the network failed to predict the true target.

Precision, recall, F1 score were calculated using the confusion matrix to evaluate the performances of the neural networks for detection of weeds growing in wheat. Precision is also called positive predictive value that measures the capability of the neural network to accurately detect the target and was calculated using the following equation:[48]

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}}$$

Recall measures the effectiveness of the neural network to correctly identify the target and was computed using the following equation:[48]

$$\text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}}$$

F1 score is the harmonic mean of precision and recall, and measures the overall performance of the neural network, which was computed using the following equation:[48]

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

3 RESULTS AND DISCUSSION

All object detection neural networks evaluated in the present study failed to show acceptable performance for detection of weed seedlings growing in wheat (Table 3). Among the object detection neural networks tested, YOLOv3 exhibited the highest F1 scores in the validation and testing datasets, mainly due to high precision (≥0.96). However, YOLOv3 exhibited unacceptably

Table 2. Hyper-parameters used for training image classification neural networks

Parameter	AlexNet	DenseNet	ResNet	VGGNet
Training epochs	90	30	30	30
Solver type	AdaDelta	SGD	Adam	AdaDelta
Batch size	2	16	16	2
Batch accumulation	5	5	5	5
Learning rate policy	Exponential decay	Lambda	Step	Exponential decay
Base learning rate	0.01	0.001	0.0001	0.01
Gamma	0.95	0.95	0.95	0.95

Table 3. Object detection neural network validation and testing results for detection of broadleaf weed seedlings in wheat

	Validation			Testing		
Neural networks	Precision	Recall	F1 score	Precision	Recall	F1 score
CenterNet2	0.68	0.58	0.62	0.59	0.51	0.55
Faster R-CNN	0.57	0.65	0.61	0.53	0.52	0.52
TridentNet	0.61	0.56	0.58	0.42	0.58	0.49
VFNet	0.54	0.62	0.58	0.48	0.52	0.50
YOLOv3	0.96	0.49	0.65	0.97	0.45	0.62

low recall values in the validation and testing datasets (≤0.49). Smaller weed targets had high detection errors, particularly false positive. A similar finding was previously noted by Sharpe *et al.*[16] in an effort to detect Carolina geranium (*Geranium carolinianum* L.) in plastic-mulched strawberry (*Fragaria* × *ananassa* Duch.) in Florida. For CenterNet, Faster R-CNN, TridentNet, and VFNet, the precision, recall, and F1 scores never exceeded 0.68. Overall, these results demonstrated that CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3 are ineffective for detecting broadleaf weed seedlings growing in wheat.

In previous studies, Dyrmann *et al.*[49,50] evaluated several other object detection neural networks for detection of weeds in winter wheat, and documented low precision and recall values. It was found that the precision and recall values of DetectNet were 0.87 and 0.46, respectively,[49] whereas that of the Single Shot MultiBox Detector were 0.82 and 0.60, respectively.[50] In other scenarios, object detection neural networks demonstrated better performance for weed detection. For example, Sharpe *et al.*[28] reported that YOLOv3 effectively detected and discriminated broadleaves, grasses, and sedges (*Cyperus* spp.) within the row-middles of vegetable plasticulture with high F1 scores (≥0.93). In bermudagrass [*Cynodon dactylon* (L.) Pers.] or perennial ryegrass (*Lolium perenne* L.) turf, DetectNet achieved high accuracy levels for weed detection.[17,18,30] For example, Yu *et al.*[17] achieved high F1 scores (≥0.98) with DetectNet for detecting dandelion (*Taraxacum officinale* Web.) in perennial ryegrass turf.

In the present study, the poor performance of weed detection with the object detection neural networks could be attributed to the following reasons: (i) unlike other row crops such as corn or cotton (*Gossypium hirsutum* L.), wheat planted in narrow row spacings (15 cm in the present study) resulted in high overlap with weeds, which increased the complexity for feature extraction, (ii) the weed species exhibit significantly different plant morphological features, which increased the computational complexity for feature extraction leading to reduced recall values, (iii) the neural networks often failed to detect weeds close to the image edges (Fig. 1), resulting in reduced recall values, although Yu *et al.*[17] suggested that continuous video inputs of smart sprayers in field applications may reduce the edge effect, and (iv) the bounding boxes generated by neural networks often covered all weeds when the testing images contained weeds at a relatively lower density (Fig. 2); nevertheless, occurrence of weed seedlings at high densities added extra complexity for feature extraction, leading to reduced recall values (Fig. 1).

The low precision values suggest that the object detection neural networks are more likely to misidentify wheat as weeds, leading to potential misapplication of herbicides. The low recall values also suggest that the object detection neural networks are more likely to misidentify weeds. This is undesirable because weeds would be missed in field applications, resulting in poor herbicide coverage and poor weed control. In previous research, Sharpe *et al.*[16,51] reported that annotation method can affect the performance of neural networks for weed detection. The authors noted that the overall accuracy of YOLOv3 for detection of goosegrass [*Eleusine indica* (L.) Gaertn.] in plastic-mulched strawberry and tomato (*Solanum lycopersicum* L.) significantly improved when the partial sections of leaf blade were annotated rather than the entire weed plant.[51] Unfortunately, it was impractical to individually annotate such small weeds in the present study, and automating annotation is a critical area for future research.

Because of the poor performance of object detection neural networks, we further explored the feasibility of using some of the image classification neural networks for detection of weed seedlings growing in wheat. It is worth noting that tp images contained weed seedlings with distinct plant morphological features and various densities of weeds growing in close proximity with wheat or even occluded under the wheat canopy, while tn images contained wheat at different densities (Fig. 3). Increasing training image numbers generally increased the F1 scores of all image classification neural networks trained with different input image sizes (Table 4). For the small training dataset, increasing training image sizes decreased the F1 scores for AlexNet and VGGNet, but increased the F1 scores for DenseNet and ResNet. However for the large training dataset, no differences were detected between the training image sizes since all neural networks exhibited remarkable performance of classification with high F1 scores (≥0.96) in the validation and testing datasets.

AlexNet and VGGNet trained with the small training dataset exhibited similar accuracy levels and achieved high F1 scores (≥0.98) in the training and testing datasets when the neural networks were trained with the input images of 200 × 200 pixels; however, these neural networks exhibited reduced F1 scores when trained with the input images of 400 × 400 pixels (Table 4). When the neural networks were trained with the small training dataset, increasing image sizes decreased the F1 scores of AlexNet and VGGNet in the validation and testing datasets, due primarily to low recall values. In the testing dataset, the precision and recall values of AlexNet were 0.91 and 0.82, respectively, whereas for VGGNet they were 0.98 and 0.85, respectively, for the image size of 400 × 400 pixels. DenseNet and ResNet showed better performances in detecting weeds growing in wheat when they were trained with large size of training images. The F1 scores of DenseNet and ResNet were ≤ 0.96 in the validation and testing datasets when they were trained with the small input image sizes of 200 × 200 pixels; however, the F1 scores increased to ≥0.98 when they were trained with a large input image size of 400 × 400 pixels.

Figure 1. YOLOv3 generated bounding boxes for detection of high density weed seedlings growing in wheat. The weed seedlings that were not detected by YOLOv3 are marked with unlabeled red bounding boxes.

Figure 2. YOLOv3 generated bounding boxes for detection of low density weed seedlings growing in wheat.

tYu *et al.*[17] evaluated the use of image classification neural networks for detection of weeds growing in bermudagrass and perennial ryegrass turf and found that VGGNet provided high overall classification accuracy when the training and testing images contained multiple broadleaf weeds with distinct plant morphological features including dandelion (*T. officinale* Web.), ground ivy (*Glechoma hederacea* L.), and spotted spurge (*Euphorbia maculata* L.). In another research, ResNet demonstrated better performance than VGGNet for weed detection in Canola (*Brassica napus* L.).[52] Further studies are needed to investigate other image classification architectures for detection of weeds in wheat.

There has been limited research evaluating the impact of image size on weed detection with DCNNs. Among the available reports, the pre-trained neural networks were optimized with specific input image sizes of 256×256 pixels for AlexNet[43] and 224×224 pixels for DenseNet,[44] ResNet,[45] and VGGNet,[46] but our results clearly showed that input image size significantly impacted classification performance. In general, AlexNet and VGGNet trained with small-sized images outperformed the

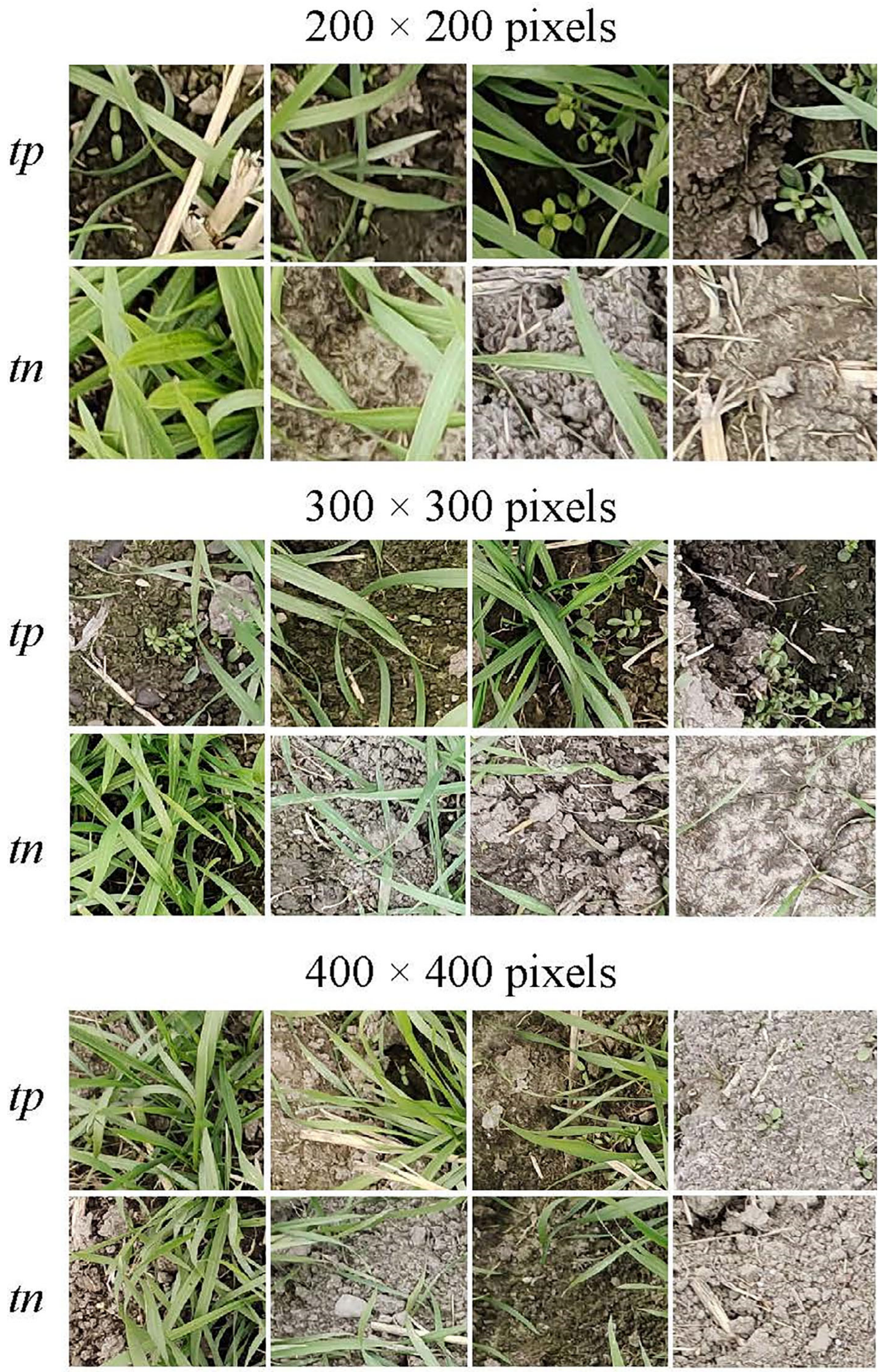

Figure 3. Images used for training image classification neural networks; tp represents true positive, tn represents true negative.

large-sized images, whereas it was the opposite for DenseNet and ResNet (large-sized images outperformed small-sized images). In this regard, He *et al.*[53] noted that the choice of image size for training DCNNs is arbitrary and training neural networks with fixed-size images may reduce recognition accuracy. Mishkin *et al.*[54] reported a similar finding that the size of training images can significantly impact recognition accuracy with DCNNs. Nevertheless, in the present study, results clearly demonstrated that increasing the number of training images can increase the precision and recall values for the neural networks investigated, thereby diminishing the effect of training image sizes.

Various herbicides, such as synthetic auxins (e.g. 2,4-D, dicamba, MCPA), sulfonylureas (e.g. florasulam, thifensulfuron, tribenuron-methyl, prosulfuron), and photosystem II inhibitors (e.g. metribuzin), are used for POST control of broadleaf weeds in wheat.[5,55–58] Autonomous robots utilizing the reliable machine vision models can significantly reduce the input of POST herbicides. It is important to note that smart sprayers

Table 4. Image classification neural networks validation and testing results for detection of broadleaf weed seedlings in wheat*

Training dataset	Neural networks		Validation			Testing		
	Deep learning architectures	Image sizes	Precision	Recall	F1 score	Precision	Recall	F1 score
Small dataset	AlexNet	200 × 200 pixels	0.99	0.99	0.99	0.98	0.99	0.98
	AlexNet	300 × 300 pixels	0.94	0.97	0.96	0.93	0.96	0.95
	AlexNet	400 × 400 pixels	0.91	0.82	0.87	0.91	0.82	0.86
	DenseNet	200 × 200 pixels	0.92	0.99	0.96	0.91	1.00	0.95
	DenseNet	300 × 300 pixels	0.94	0.98	0.96	0.88	1.00	0.94
	DenseNet	400 × 400 pixels	0.98	1.00	0.99	0.98	0.98	0.98
	ResNet	200 × 200 pixels	0.92	1.00	0.96	0.90	1.00	0.95
	ResNet	300 × 300 pixels	0.95	0.97	0.96	0.91	1.00	0.96
	ResNet	400 × 400 pixels	1.00	0.99	1.00	1.00	0.98	0.99
	VGGNet	200 × 200 pixels	1.00	0.99	0.99	0.99	0.99	0.99
	VGGNet	300 × 300 pixels	0.99	0.98	0.98	0.98	0.98	0.98
	VGGNet	400 × 400 pixels	0.94	0.93	0.94	0.98	0.85	0.91
Large dataset	AlexNet	200 × 200 pixels	1.00	1.00	1.00	0.99	0.99	0.99
	AlexNet	300 × 300 pixels	0.97	0.99	0.98	0.95	0.98	0.96
	AlexNet	400 × 400 pixels	1.00	0.98	0.99	1.00	0.94	0.97
	DenseNet	200 × 200 pixels	1.00	1.00	1.00	1.00	1.00	1.00
	DenseNet	300 × 300 pixels	0.96	1.00	0.98	0.95	1.00	0.97
	DenseNet	400 × 400 pixels	1.00	1.00	1.00	0.98	0.97	0.98
	ResNet	200 × 200 pixels	1.00	1.00	1.00	0.99	1.00	1.00
	ResNet	300 × 300 pixels	0.98	0.99	0.99	0.96	1.00	0.98
	ResNet	400 × 400 pixels	1.00	1.00	1.00	1.00	0.97	0.99
	VGGNet	200 × 200 pixels	1.00	1.00	1.00	1.00	0.99	1.00
	VGGNet	300 × 300 pixels	0.98	0.99	0.99	0.99	1.00	0.99
	VGGNet	400 × 400 pixels	1.00	1.00	1.00	0.99	0.99	0.99

*The neural networks were developed by training AlexNet and VGGNet with input images of 200 × 200, 300 × 300, or 400 × 400 pixels.

relying on image classification neural networks as the decision system cannot utilize nozzles with narrow spray patterns to target individual weeds. Nevertheless, the image classification neural networks trained with small size images may allow the smart sprayer to deliver herbicides to the small area containing weeds, thus saving more herbicides compared to the networks trained with large size images.

4 CONCLUSIONS

The evaluated object detection neural networks including CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3 failed to reliably detect broadleaf weed seedlings growing in wheat; the precision and recall values in both validation and testing datasets were low, but this issue was overcome by image classification neural networks. When the neural networks were trained with the small training dataset, increased classification accuracy (high recall values) was observed with AlexNet and VGGNet when they were trained with the images of 200 × 200 pixels than 300 × 300 or 400 × 400 pixels. Conversely, decreased classification accuracy (reduced precision values) was noted for DenseNet and ResNet when they were trained with the images of 200 × 200 pixels, rather than 300 × 300 or 400 × 400 pixels. Nevertheless, increasing the number of training images improved the performance of classification for all neural networks, regardless of training image sizes. When the neural networks were trained with the large training dataset and small size training images of 200 × 200 pixels, all image classification neural networks demonstrated remarkable performance at detecting weed seedlings in wheat, with high F1 score values (≥0.99). To further improve the performance of weed detection in wheat, other DCNNs architectures may be investigated in the future.

ACKNOWLEDGEMENTS

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

CONFLICT OF INTEREST

No conflicts of interest have been declared.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1 Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M and Muricho G, Crops that feed the world. Past successes and future challenges to the role played by wheat in global food security. *Food Secur* **5**:291–317 (2013).
2 Wells G, Annual weed competition in wheat crops: the effect of weed density and applied nitrogen. *Weed Res* **19**:185–191 (1979).
3 Melander B and McCullough M, Influence of intra-row cruciferous surrogate weed growth on crop yield in organic spring cereals. *Weed Res* **60**:464–474 (2020).
4 Wang H, Liu W, Jin T, Peng X, Zhang L and Wang J, Bipyrazone: a new HPPD-inhibiting herbicide in wheat. *Sci Rep* **10**:1–10 (2020).

5 Zargar M, Bayat M and Astarkhanova T, Study of postemergence-directed herbicides for redroot pigweed (*Amaranthus retroflexus* L.) control in winter wheat in southern Russia. *J Plant Prot Res* **60**:7–13 (2020).
6 Heap I, *The International Herbicide-Resistant Weed Database*. Available: www.weedscience.org [2 May 2021].
7 Fennimore SA and Cutulle M, Robotic weeders can improve weed control options for specialty crops. *Pest Manag Sci* **75**:1767–1774 (2019).
8 Wang A, Zhang W and Wei X, A review on weed detection using ground-based machine vision and image processing techniques. *Comp Electron Agric* **158**:226–240 (2019).
9 Longchamps L, Panneton B, Samson G, Leroux GD and Thériault R, Discrimination of corn, grasses and dicot weeds by their UV-induced fluorescence spectral signature. *Precis Agric* **11**:181–197 (2010).
10 Peteinatos GG, Weis M, Andújar D, Rueda Ayala V and Gerhards R, Potential use of ground-based sensor technologies for weed detection. *Pest Manag Sci* **70**:190–199 (2014).
11 Su WH, Fennimore SA and Slaughter DC, Fluorescence imaging for rapid monitoring of translocation behaviour of systemic markers in snap beans for automated crop/weed discrimination. *Biosyst Eng* **186**:156–167 (2019).
12 Liang W, Ma X, Wan P and Liu L, Plant salt-tolerance mechanism: a review. *Biochem Biophy Res Commun* **495**:286–291 (2018).
13 Wu D, Feng L, Zhang C and He Y, Early detection of *Botrytis cinerea* on eggplant leaves based on visible and near-infrared spectroscopy. *Trans ASABE* **51**:1133–1139 (2008).
14 Feyaerts F and Van Gool L, Multi-spectral vision system for weed detection. *Pattern Recogit Lett* **22**:667–674 (2001).
15 Sapkota B, Singh V, Neely C, Rajan N and Bagavathiannan M, Detection of Italian ryegrass in wheat and prediction of competitive interactions using remote-sensing and machine-learning techniques. *Remote Sens* **12**:2977 (2020).
16 Sharpe SM, Schumann AW and Boyd NS, Detection of Carolina geranium (*Geranium carolinianum*) growing in competition with strawberry using convolutional neural networks. *Weed Sci* **67**:239–245 (2018).
17 Yu J, Schumann AW, Cao Z, Sharpe SM and Boyd NS, Weed detection in perennial ryegrass with deep learning convolutional neural network. *Front Plant Sci* **10**:1422 (2019).
18 Yu J, Sharpe SM, Schumann AW and Boyd NS, Deep learning for image-based weed detection in turfgrass. *Eur J Agron* **104**:78–84 (2019).
19 LeCun Y, Bengio Y and Hinton G, Deep learning. *Nature* **521**:436–444 (2015).
20 Guo Y, Liu Y, Oerlemans A, Lao S, Wu S and Lew MS, Deep learning for visual understanding: a review. *Neurocomputing* **187**:27–48 (2016).
21 Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K and Kuksa P, Natural language processing (almost) from scratch. *J Mach Learn Res* **12**:2493–2537 (2011).
22 Gawehn E, Hiss JA and Schneider G, Deep learning in drug discovery. *Mol Inform* **35**:3–14 (2016).
23 Khaki S and Wang L, Crop yield prediction using deep neural networks. *Front Plant Sci* **10**:621 (2019).
24 Kang H and Chen C, Fast implementation of real-time fruit detection in apple orchards using deep learning. *Comput Electron Agric* **168**: 105108 (2020).
25 Ghosal S, Blystone D, Singh AK, Ganapathysubramanian B, Singh A and Sarkar S, An explainable deep machine vision framework for plant stress phenotyping. *Proc Natl Acad Sci USA* **115**:4613–4618 (2018).
26 Peteinatos G, Reichel P, Karouta J, Andújar D and Gerhards R, Weed identification in maize, sunflower, and potatoes with the aid of convolutional neural networks. *Remote Sens* **12**:4185 (2020).
27 Sabzi S, Abbaspour-Gilandeh Y and García-Mateos G, A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms. *Comput Ind* **98**:80–89 (2018).
28 Sharpe SM, Schumann AW, Yu J and Boyd NS, Vegetation detection and discrimination within vegetable plasticulture row-middles using a convolutional neural network. *Precis Agric* **21**:264–277 (2019).
29 dos Santos FA, Freitas DM, da Silva GG, Pistori H and Folhes MT, Weed detection in soybean crops using ConvNets. *Comput Electron Agric* **143**:314–324 (2017).
30 Yu J, Schumann AW, Sharpe SM, Li X and Boyd N, Detection of grassy weeds in bermudagrass with deep convolutional neural networks. *Weed Sci* **68**:545–552 (2020).
31 Su D, Qiao Y, Kong H and Sukkarieh S, Real time detection of inter-row ryegrass in wheat farms using deep learning. *Biosyst Eng* **204**:198–211 (2021).
32 Yu J and McCullough PE, Growth stage influences mesotrione efficacy and fate in two bluegrass (*Poa*) species. *Weed Technol* **30**:524–532 (2016).
33 Kieloch R and Domaradzki K, The role of the growth stage of weeds in their response to reduced herbicide doses. *Acta Agrobot* **64**:259–266 (2011).
34 Zhou X, Wang D and Krähenbühl P, Objects as points. arXiv:1904.07850 (2019).
35 Ren S, He K, Girshick R and Sun J, eFaster R-CNN: towards real-time object detection with region proposal networks. arXiv:150601497 (2015).
36 Li Y, Chen Y, Wang N and Zhang Z, Scale-aware trident networks for object detection. arXiv:1901.01892 (2019).
37 Ahmed A, Tangri P, Panda A, Ramani D and Nevronas S, VFNet: a convolutional architecture for accent classification. arXiv:1910.06697 (2019).
38 Redmon J and Farhadi A, Yolov3: an incremental improvement. arXiv: 180402767 (2018).
39 Redmon J, Divvala S, Girshick R and Farhadi A, You only look once: unified, real-time object detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); June 27–30, 2016; Las Vegas, NV, USA*. IEEE (2016).
40 Redmon J and Farhadi A, YOLO9000: better, faster, stronger, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); July 21–26, 2017; Honolulu, HI, USA*. IEEE (2017).
41 Sa I, Ge Z, Dayoub F, Upcroft B, Perez T and McCool C, Deepfruits: a fruit detection system using deep neural networks. *Sensors* **16**:1222 (2016).
42 Lin TY, Maire M, Belongie S, Hays J and Zitnick CL, Microsoft COCO: common objects in context, in lecture notes in computer science, in *Proceedings of the 13th European Conference on Computer Vision; September 6-12, 2014; Zurich, Switzerland*. Springer, New York (2014).
43 Krizhevsky A, Sutskever I and Hinton GE, Imagenet classification with deep convolutional neural networks. *Commun ACM* **60**:84–90 (2012).
44 Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T and Keutzer K., Densenet: implementing efficient convnet descriptor pyramids. arXiv:1404.1869 (2014).
45 He K, Zhang X, Ren S and Sun J, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); June 27–30, 2016; Las Vegas, NV, USA*. IEEE (2016).
46 Simonyan K and Zisserman A, Very deep convolutional networks for large-scale image recognition, in *Proceedings of the 3rd International Conference on Learning Representations, ICLR; May 7-9, 2015; San Diego, CA, USA*. ICLR (2015).
47 Deng J, Dong W, Socher R, Li L, Li K, Li F, Imagenet: a large-scale hierarchical image database, in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition; June 20–25, 2009; Miami, FL, USA*. IEEE (2009).
48 Sokolova M and Lapalme G, A systematic analysis of performance measures for classification tasks. *Inf Process Manag* **45**:427–437 (2009).
49 Dyrmann M, Jørgensen RN and Midtiby HS, RoboWeedSupport-detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network. *Adv Anim Biosci* **8**:842–847 (2017).
50 Dyrmann M, Skovsen S, Laursen MS and Jørgensen RN, Using a fully convolutional neural network for detecting locations of weeds in images from cereal fields, in *Proceedings of the 14th International Conference on Precision Agriculture; June 24–27, 2018; Montreal, Quebec, Canada*. International Society of Precision Agriculture, Monticello (2018).
51 Sharpe SM, Schumann AW and Boyd NS, Goosegrass detection in strawberry and tomato using a convolutional neural network. *Sci Rep* **10**:9548 (2020).

52 Asad MH and Bais A, Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. *Inf Process Agric* **7**:535–545 (2019).
53 He K, Zhang X, Ren S and Sun J, Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* **37**:1904–1916 (2015).
54 Mishkin D, Sergievskiy N and Matas J, Systematic evaluation of convolution neural network advances on the imagenet. *Comput Vis Image Underst* **161**:11–19 (2017).
55 Baghestani MA, Zand E, Soufizadeh S, Beheshtian M, Haghighi A, Barjasteh A *et al*., Study on the efficacy of weed control in wheat (*Triticum aestivum* L.) with tank mixtures of grass herbicides with broadleaved herbicides. *Crop Prot* **27**:104–111 (2008).
56 Chhokar R, Sharma R, Chauhan D and Mongia A, Evaluation of herbicides against *Phalaris minor* in wheat in north-western Indian plains. *Weed Res* **46**:40–49 (2006).
57 Curran WS, Wallace JM, Mirsky S and Crockett B, Effectiveness of herbicides for control of hairy vetch (*Vicia villosa*) in winter wheat. *Weed Technol* **29**:509–518 (2015).
58 Zargar M, Bayat M, Lyashko M and Chauhan B, Postemergence herbicide applications impact Canada thistle control and spring wheat yields. *Agron J* **111**:2874–2880 (2019).

