

Article

Detection of Weeds Growing in Alfalfa Using Convolutional Neural Networks

Jie Yang ^{1,2} , Yundi Wang ³, Yong Chen ^{1,*} and Jialin Yu ^{2,*}

¹ College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China; yangjiehahay@163.com

² Peking University Institute of Advanced Agricultural Sciences, Weifang 261325, China

³ Department of Computer Science, Stevens Institute of Technology, Hoboken, NJ 07030, USA; wyundi8615@outlook.com

* Correspondence: chenrongjsnj@163.com (Y.C.); jialin.yu@pku-iaas.edu.cn (J.Y.)

Abstract: Alfalfa (*Medicago sativa* L.) is used as a high-nutrient feed for animals. Weeds are a significant challenge that affects alfalfa production. Although weeds are unevenly distributed, herbicides are broadcast-applied in alfalfa fields. In this research, object detection convolutional neural networks, including Faster R-CNN, VarifocalNet (VFNet), and You Only Look Once Version 3 (YOLOv3), were used to indiscriminately detect all weed species (1-class) and discriminately detect between broadleaves and grasses (2-class). YOLOv3 outperformed other object detection networks in detecting grass weeds. The performances of using image classification networks (GoogLeNet and VGGNet) and object detection networks (Faster R-CNN and YOLOv3) for detecting broadleaves and grasses were compared. GoogLeNet and VGGNet (F_1 scores ≥ 0.98) outperformed Faster R-CNN and YOLOv3 (F_1 scores ≤ 0.92). Classifying and training various broadleaf and grass weeds did not improve the performance of the neural networks for weed detection. VGGNet was the most effective neural network (F_1 scores ≥ 0.99) tested to detect broadleaf and grass weeds growing in alfalfa. Future research will integrate the VGGNet into the machine vision subsystem of smart sprayers for site-specific herbicide applications.

Keywords: alfalfa; deep convolutional neural networks; object detection; image classification; precision herbicide application



Citation: Yang, J.; Wang, Y.; Chen, Y.; Yu, J. Detection of Weeds Growing in Alfalfa Using Convolutional Neural Networks. *Agronomy* **2022**, *12*, 1459. <https://doi.org/10.3390/agronomy12061459>

Academic Editors: Sudheesh Manalil and Anestis Karkanis

Received: 31 May 2022

Accepted: 16 June 2022

Published: 17 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Alfalfa (*Medicago sativa* L.), a perennial crop, is the preferred feed for livestock such as dairy cows [1]. Alfalfa is rich in protein, vitamins, minerals, and fiber [2], making the crop high nutritional forage [3]. Weeds are a significant challenge in alfalfa production. They compete with alfalfa for nutrients, space, sunlight, and water, reducing forage quality and yield. In addition, some weed species, such as perilla mint (*Perilla frutescens* L.), contain substances that are toxic to livestock [4]. A variety of postemergence (POST) herbicides are broadcast-applied for weed control in alfalfa, although weeds are almost always unevenly distributed in fields. For example, clethodim and 2,4-DB control a wide range of grasses and broadleaves in conventional alfalfa, respectively [5,6], while glyphosate provides nonselective control of weeds in glyphosate-tolerant alfalfa [7]. A machine vision-based smart sprayer may precisely apply herbicides to the broadleaf and grass weeds in alfalfa, thus reducing herbicide input. Compared to a broadcast herbicide application, targeted and precise spraying at specific sites may reduce herbicide inputs by 90% [8].

Site-specific weed management, particularly precision herbicide application, can considerably reduce herbicide input and weed control costs [9–12]. The major obstacle to perform autonomous precision herbicide application is the accurate and reliable detection of weeds in real-time. Traditional machine vision techniques are based on plant leaf color, spectral information, feature fusion [10,13], morphological features [14–16], and spatial

information [17]. However, these approaches cannot reliably detect weeds intermingled with crops, especially in a complex environment with high crop and weed densities [18–21].

In recent years, machine learning techniques have advanced rapidly [22]. Deep convolutional neural networks (DCNNs) have succeeded in various applications [23,24]. For example, recent studies have shown that deep learning can be used to diagnose coronavirus disease [25], help predict seizure recurrence [26], perform high-accuracy three-dimensional optical measurement [27], predict the activity of potential drug molecules [28], analyze particle accelerator data [29,30], defect detection of industry wood veneer [31], achieve efficient classification of green plum detects [32], and reconstruct brain circuits [33]. In addition, DCNNs have demonstrated an exceptional capability for object detection [34,35] and classification in digital images [36,37].

Researchers have explored the feasibility of using DCNNs for weed detection in various cropping systems, such as bermudagrass and vegetable fields [11,38–42]. Sharpe et al. [42] showed that YOLOv3 could be used as an object detector to discriminate broadleaves, grasses, and sedges in the row middles of plastic-mulched vegetable crops. Yu et al. [11,21] reported the feasibility of using DCNNs for the detection of multiple broadleaf and grass weeds in actively growing or dormant bermudagrass (*Cynodon dactylon* (L.) Pers.). Hennessy et al. [43] reported the feasibility of using YOLO3-tiny to detect hairy fescue (*Festuca filiformis* Pourr.) and sheep sorrel (*Rumex acetosella* L.) in wild blueberry (*Vaccinium* spp. L.). Hussain et al. [44] investigated the feasibility of using DCNNs for detecting common lambsquarters (*Chenopodium album* L.) in potato (*Solanum tuberosum* L.). However, the feasibility and effectiveness of utilizing DCNNs for weed detection in alfalfa has never been investigated.

Alfalfa hay is typically harvested multiple times per growing season, unlike previously mentioned crops. Alfalfa can re-grow following harvest and can rapidly regenerate new stems and leaves. Weed detection in various heights of alfalfa stands might be a significant challenge. The objective of the research was to evaluate the use of DCNNs for detecting weeds growing in alfalfa.

2. Materials and Methods

2.1. Overview

The five DCNNs, GoogLeNet [45], VGGNet [46], Faster R-CNN [47], VarifocalNet (VFNet) [48], and YOLOv3 [49], were evaluated for detection of weeds growing in alfalfa. These neural networks were pre-trained with 256×256 pixels images and employed the ADADELTA deep learning optimizer for backpropagation [50]. This method is used to overcome the fundamental problem of deep learning—the activation function signal error in the progress of cumulative backpropagation increases or decreases rapidly [51,52]. GoogLeNet and VGGNet are convolutional neural networks for image classification [45,46]. The image classification networks mainly achieve the purpose of classifying or predicting object categories by accurately detecting the objects in images. GoogLeNet consists of 22 convolutional layers and is designed on small convolutions in order to reduce the neuron numbers and parameters [45]. VGGNet used in this research is composed of 19 weight layers. VGGNet is designed to implement smaller convolutional kernels to limit neuron numbers and parameters [46]. Faster R-CNN, VarifocalNet (VFNet), and YOLOv3 are convolutional neural networks for object detection [47–49]. In addition to the purpose of classification, the object detection networks can also accurately detect the location of the target in images and mark the corresponding label. In terms of structure, Faster R-CNN has integrated feature extraction, proposal extraction, bounding box regression (rect refine), and classification into one network, which significantly improves the overall performance, especially in terms of detection speed [47]. VFNet is proposed to learn the intersection over union (IoU)-aware classification score (IACS), which can simultaneously represent the confidence of the object's existence and the positioning accuracy to achieve a more accurate detection in the dense object detector [48]. At the same time, a new function called Varifocal loss was designed to train dense object detectors to predict IACS, and a

new efficient star-shaped bounding box feature representation was designed to estimate IACS and improve rough bounding boxes. YOLOv3 is a fully convolutional network that uses many residual layer jump connections [49]. In order to reduce the negative effect of gradient caused by pooling, the author directly abandoned POOLing and used conv's stride to achieve downsampling. YOLOv3 uses upsample and fusion methods similar to Feature Pyramid Network (FPN) to detect multiple-scale feature maps to enhance the algorithm's accuracy for small target detection. Different from Faster R-CNN, YOLOv3 only operates on a single best priority. All image classification neural networks were pre-trained using the ImageNet database [53], with specific spatial tensor image sizes of 224×224 pixels [45,46,54,55].

2.2. Image Acquisition

Images of various weed species growing in alfalfa were acquired multiple times during September and October 2020 using a digital camera (Panasonic® DMC-ZS110, Xiamen, Fujian, China) at a resolution of 4160×3120 pixels. The images taken in alfalfa fields in Bengbu, Anhui, China ($117^{\circ}89' \text{ N}$, $117^{\circ}88' \text{ E}$), were used for the training, validation, and testing datasets. Additional images were taken in separate alfalfa fields in Bengbu, Anhui, China, and Yangzhou University Pratacultural Science Experiment Station in Yangzhou, Jiangsu, China ($32^{\circ}20' \text{ N}$, $119^{\circ}23' \text{ E}$). The images containing alfalfa (measured 8 to 52 cm in height) and various broadleaf and grass weed species were captured from approximately 1.5 m from the ground level, yielding $0.05 \text{ cm pixel}^{-1}$. The images were acquired under various outdoor lighting conditions, including clear/bright, cloudy, or partly cloudy skies.

2.3. Training and Testing

The training or testing dataset contained a variety of broadleaf and grass weed species occurring in the mixture. The dominant broadleaf weed species were Carolina geranium (*Geranium carolinianum* L.), catchweed bedstraw (*Galium apaine* L.), mugwort (*Artemisia Vulgaris* L.), and speedwell (*Veronica* spp. L.), whereas the dominant grass weeds were barnyardgrass (*Echinochloa crus-galli* (L.) Beauv), crabgrass (*Digitaria* spp.), and goosegrass (*Eleusine indica* (L.) Gaertn).

Four sets of labels were produced for object detection of convolutional neural networks, one per network. Images were resized to 1280×720 pixels (720 p) using Irfanview (Version 4.50, Irfan Skijan, Bosnia). The resolution was chosen to integrate the neural networks into developed precision sprayer technology using 720p video as input images. The object weeds were labeled with labeling (<https://github.com/tzutalin/labelImg>, accessed on 9 January 2021). The following neural networks were trained to detect weeds:

The neural networks trained to exclusively detect broadleaf weeds were termed 1 class (B). A total of 573 images containing broadleaves were included in the training datasets. All broadleaf weeds were labeled under a single category.

The neural networks trained to exclusively detect grass weeds were termed 1 class (G). A total of 926 images containing grasses were used for the training dataset. All grass weeds were labeled under a single category.

The neural networks trained to indiscriminately detect both broadleaf and grass weeds were termed 1 class (B + G) (broadleaves + grasses). All broadleaf and grass weeds were labeled under a single category. A total of 873 images were used for the training dataset. All weeds were labeled under a single category.

The neural networks were trained to detect and discriminate between broadleaf and grass weeds were termed 2 classes (B + G) (broadleaves + grasses). A total of 935 images containing 291 broadleaves and 6495 grasses were used for the training dataset. The images collected in Bengbu, Anhui Province, were selected as the training dataset. The predominant weed species was grasses with few broadleaf weeds. Weed species were labeled under separate categories based on the herbicide weed control spectrum described above. All annotations were entirely composed of a single bounding box.

The images used for image classification were divided into equal sub-images with 256×256 pixels through cropping and then classified according to their species. The weed species were roughly divided into broadleaf (positive) and grass (negative) weeds. The training dataset contained 1700 positive and 1700 negative images. The validation dataset contained 340 positive and 340 negative images. The testing dataset contained 170 positive and 170 negative images. The multiple-species neural networks were trained using a dataset containing 4000 positive images (1000 images for each broadleaf weeds) and 9000 negative images (images without targeted broadleaf weeds). The validation dataset contained 200 positive and 200 negative images, and the testing dataset contained 100 positive and 100 negative images for each weed species.

Image classification architectures, including GoogLeNet and VGGNet, were trained for weed detection. Data were imported into the NVIDIA Deep Learning GPU Training System (DIGITS) (Version 6.0.0, NVIDIA, Santa Clara, CA, USA). The training and testing were performed on a GeForce RTX 2080Ti with 64 GB of memory using the Convolutional Architecture for Fast Feature Embedding (CAFFE) [56]. The hyper-parameters used for training the neural networks are presented in Table 1.

Table 1. Hyper-parameters used for training the neural networks.

	GoogLeNet and VGGNet	Faster R-CNN and VNet	YOLOV3
Training epochs	30	100	273
Solver type	SGD	SGD	SGD
Batch size	2	- ^a	64
Batch accumulation	5	-	-
Learning rate policy	Step Down	Step Down	Step Down
Base learning rate	0.01	0.001	0.001
Gamma	0.1	2.0	2.0
Step Size	33%	-	-

Abbreviation: SGD, Stochastic Gradient Descent. ^a The hyphen “-” indicates that there is no fixed value for this item.

Object detection uses input images and label files containing the targeted object in each image. The neural networks selected were Faster-R CNN, VNet, and YOLOv3. Faster-R CNN and VNet were trained and tested using mmDetection neural network framework [57] and pretrained using Pattern Analysis, Statistical Modeling, and Computational Learning Visual Object Classes (PASCAL VOC) dataset. YOLOv3 was trained and tested using the Darknet neural network framework [58] and pretrained using the PASCAL VOC dataset. The hyper-parameters used for training the neural networks are presented in Table 1. The training and validation relied on the union intersection between ground truth labels and predicted bounding boxes. Therefore, based on the two standards related to the network application of precise spraying, the validation results were evaluated using $IoU > 0$ to visually identify the actual vegetation in the image.

The validation and testing results of the neural networks were arranged in a confusion matrix with four possible conditions: true positive (*tp*), false positive (*fp*), false negative (*fn*), and true negative (*tn*). A *tp* is when the neural network correctly identifies the target. A *fp* is when the neural network falsely identifies a target. A *fn* is when the neural network fails to identify a target. Although true negative (*tn*) does complete the confusion matrix, this category is not a key priority for current applications. Precision, recall, and F_1 score were computed based on the results of confusion matrices.

Precision measures the accuracy of the neural network at positive detection and was calculated with Equation (1) [59–61]:

$$\text{Precision} = \frac{tp}{tp + fp} \quad (1)$$

Recall measures the effectiveness of the neural network in identifying the target and was determined using the Equation (2) [59–61]:

$$\text{Recall} = \frac{tp}{tp + fn} \quad (2)$$

F₁ score is the harmonic mean of precision and recall. The F₁ score is used for comprehensive evaluation of precision and recall and was calculated using Equation (3) [60]:

$$F_1 \text{ score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (3)$$

3. Results

3.1. Object Detection

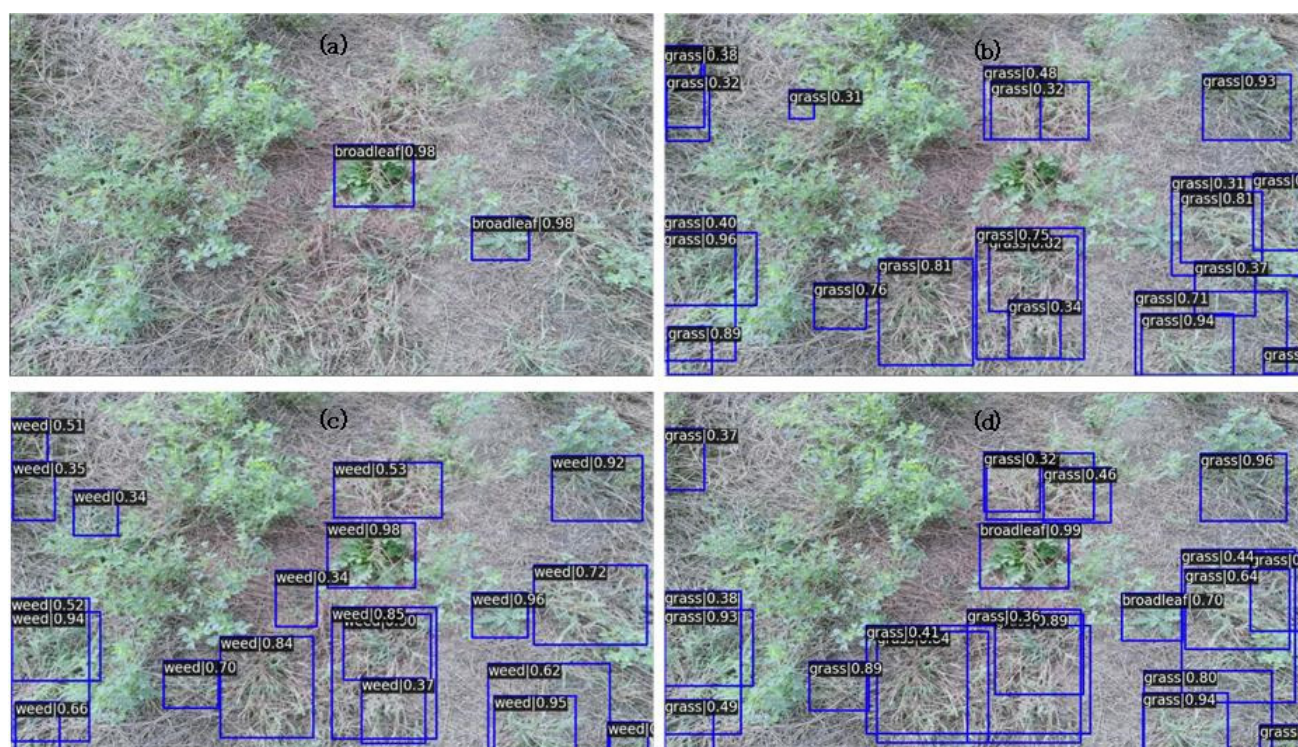
All object detection neural networks, including the 1-class (B), 1-class (G), 1-class (B + G), and 2-class (B + G), showed unacceptable performance of weed detection, as evidenced by low precision, recall, and F₁ score values (Table 2). To detect broadleaf weeds growing in alfalfa, the F₁ scores of 1-class (B) never exceeded 0.27, 0.69, and 0.17 for Faster R-CNN, YOLOv3, and VFNet, respectively. The effect diagram was shown in Figure 1. The low precision and recall values of 1-class (B) are likely due to the fact that broadleaf weeds and alfalfa are dicotyledons and share similar plant morphological characteristics, increasing the difficulty of feature extraction and resulting in poor performance of weed detection. For detection of grass weeds growing in alfalfa, the 1-class (G) trained with YOLOv3 outperformed Faster R-CNN and VFNet, but the F₁ score never exceeded 0.91.

Table 2. Object detection neural network validation results for detection of weeds growing in alfalfa ^a.

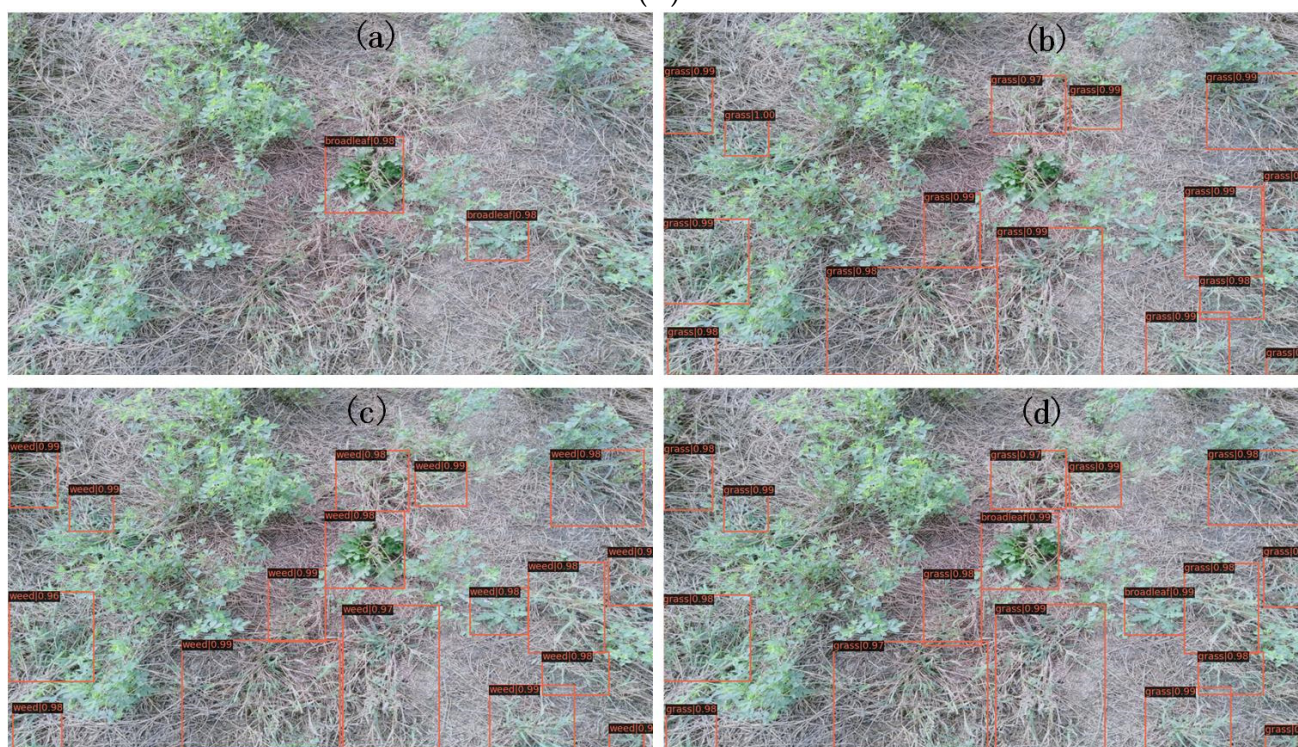
Model	Class	Precision	Recall	F1 Score
Faster R-CNN	1-class (B)	0.21	0.38	0.27
	1-class (G)	0.55	0.82	0.66
	1-class (B + G)	0.47	0.72	0.57
	2-class	0.53	0.66	0.59
VFNet	1-class (B)	0.47	0.11	0.17
	1-class (G)	0.68	0.53	0.59
	1-class (B + G)	0.71	0.54	0.62
	2-class	0.67	0.56	0.61
YOLO v3	1-class (B)	0.91	0.55	0.69
	1-class (G)	0.96	0.87	0.91
	1-class (B + G)	0.92	0.60	0.73
	2-class	0.62	0.84	0.71

^a 1-class (B) refers to the training dataset only containing broadleaf weeds and the neural network trained to exclusively detect broadleaf weeds; 1-class (G) refers to the training dataset only containing grass weeds and the neural network trained to exclusively detect grass weeds; 1-class (B + G) refers to the training dataset contained both broadleaf and grass weeds and the neural network trained to indiscriminately detect both broadleaf and grass weeds; 2-class refers to the training dataset contained broadleaf and grass weeds and the neural network trained to discriminate between broadleaf and grass weeds.

The 1-class (G) and 1-class (B + G) neural networks generally outperformed the 1-class (B) and 2-class (B + G) neural networks in detecting the target weeds (Table 2). Among the evaluated neural networks, the 1-class (B) neural network exhibited the worst weed detection, and the F₁ scores of Faster R-CNN, VFNet, or YOLOv3 did not exceed 0.69. The 1-class (G) trained with YOLOv3 exhibited the highest F₁ score (0.91). To detect broadleaf and grass weeds growing in alfalfa, the 1-class (B + G) trained with YOLOv3 showed the highest F₁ score (0.73). To detect and discriminate broadleaf and grass weeds growing in alfalfa, the 2-class neural network trained with YOLOv3 showed the highest F₁ score (0.71). The major weed species in the training images were grasses, with few broadleaves. This may have resulted in insufficient training samples of broadleaves, leading to the low precision and recall values of 1-class (B). An additional study is needed to train 1-class (B) using the training images containing more broadleaf weeds.



(A)



(B)

Figure 1. Cont.

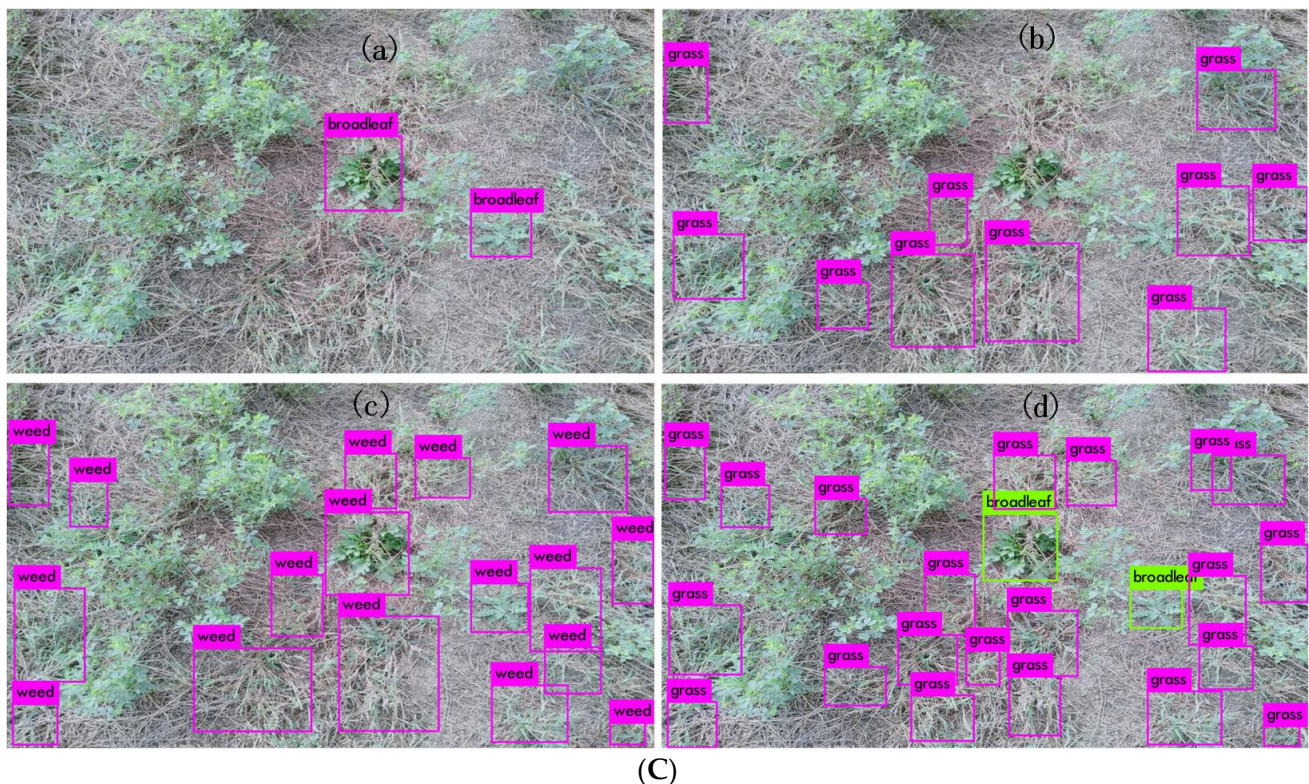


Figure 1. The detection of weeds growing in alfalfa using object detection neural networks. (A) Weed detection based on Faster R-CNN. (B) Weed detection based on VFNet. (C) Weed detection based on YOLOv3. (a) The training dataset only contained broadleaf weeds to detect broadleaf weeds growing in alfalfa. (b) The training dataset only contained grass weeds to detect grasses growing in alfalfa. (c) The training dataset contained both broadleaf and grass weeds to indiscriminately detect both broadleaf and grass weeds growing in alfalfa. (d) The training dataset contained broadleaf and grass weeds to detect and discriminate between broadleaf and grass weeds growing in alfalfa.

3.2. Image Classification vs. Object Detection

The GoogLeNet and VGGNet outperformed Faster R-CNN and YOLOv3 in validation and testing results to detect broadleaf and grass weeds growing in alfalfa (Table 3). To detect broadleaf or grass weeds, the F_1 scores of GoogLeNet and VGGNet networks were ≥ 0.98 . In contrast, for detecting the same target weeds, the F_1 scores of Faster R-CNN and YOLOv3 never exceeded 0.92. The excellent performances of image classification neural networks for weed detection can likely be attributed to the following reasons: (1) the cropped images may enlarge the features of the positive area to a certain extent, while the negative area may accordingly be reduced; (2) because of the selected images were cropped, the number of training images increased; (3) the pixels to be analyzed in the cropped images were relatively few. These facts may have increased the percentage of images with positive pattern recognition, which was amplified by the convolutional filters, improving the performance of weed detection.

In the validation results, the precision values of broadleaf and grass weeds trained by the YOLOv3 reached 1.00 and 0.99, while the recall was ≤ 0.52 . In the testing results, the precision and recall values of the grass weeds were 0.89 and 0.95. The precision of the broadleaves was 0.97, while the recall was 0.67. The low recall was due to the fact that some samples were not correctly detected, and the threshold for predicting true positive samples of the two-classifier was decreased.

Table 3. The validation and testing results using image classification and object detection neural networks for the detection of weeds growing in alfalfa ^b.

Model	Weed Species	Network Type ^a	Validation Results			Testing Results		
			Precision	Recall	F ₁ Score	Precision	Recall	F ₁ Score
GoogLeNet	Broadleaves	IC	0.99	0.98	0.99	0.98	0.97	0.98
	Grasses	IC	0.99	0.98	0.98	0.98	0.98	0.98
VGGNet	Broadleaves	IC	0.99	0.99	0.99	1.00	0.98	0.99
	Grasses	IC	0.99	0.99	0.99	0.98	1.00	0.99
Faster R-CNN	Broadleaves	OD	0.22	0.32	0.26	0.23	0.46	0.31
	Grass	OD	0.68	0.40	0.50	0.87	0.52	0.65
YOLOv3	Broadleaves	OD	1.00	0.46	0.63	0.97	0.67	0.79
	Grasses	OD	0.99	0.52	0.68	0.89	0.95	0.92

^a Abbreviations: IC, image classification; OD, object detection. ^b The models were trained with the training dataset containing various broadleaf and grass weeds. The image classification of the training dataset contained 3000 positive and 3000 negative images. The validation dataset contained 600 positive and 600 negative images. The testing dataset contained 300 positive and 300 negative images. The object detection neural networks for detecting grass weeds contained 926 images, while the object detection neural networks for detecting broadleaf weeds contained 532 images. The validation dataset contained 93 images, while the testing dataset contained 100 images.

3.3. Broadleaf and Grass Weeds Detection Using Various Convolutional Neural Networks

VGGNet outperformed GoogLeNet, Faster R-CNN, and YOLOv3 to detect broadleaves, including Carolina geranium, catchweed bedstraw, mugwort, and speedwell growing in alfalfa (Table 4). For detection of speedwell, the precision, recall, and F₁ score values of VGGNet reached 1.00. For the image classification neural networks, the F₁ scores of GoogLeNet for detecting speedwell were ≥ 0.87 , while the F₁ scores of GoogLeNet for detecting these four species of broadleaves never exceeded 0.89. Meanwhile, the F₁ scores of VGGNet for detection of speedwell reached 1.00, while the F₁ scores of VGGNet for detection of these broadleaf weed species were ≥ 0.92 . These findings suggest that these two image classification neural networks effectively detected speedwell. For the object detection neural networks, YOLOv3 outperformed Faster R-CNN. For detection of mugwort, the F₁ scores of Faster R-CNN never exceeded 0.57, while the F₁ scores of YOLOv3 reached 0.84. To detect four broadleaf weed species, the F₁ score of YOLOv3 ranged from 0.15 to 0.84. In addition, the F₁ score of Faster R-CNN never exceeds 0.57. These two object detection neural networks showed an unacceptable performance for detecting these four species of broadleaf weeds.

To detect grass weeds, VGGNet outperformed GoogLeNet, Faster R-CNN, and YOLOv3 (Table 5). The image classification networks, GoogLeNet and VGGNet, showed better weed detection performances than the object detection networks, Faster R-CNN and YOLOv3, and showed lower F₁ scores when detecting broadleaf weeds than grass weeds. Image segmentation may be an optional approach for such circumstances [62]. Compared with Section 3.2, specific detection of weed species did not improve the performance of networks. For example, the F₁ scores of VGGNet for detection of grasses did not exceed 0.76, which did not meet the standards for precision spraying. However, the F₁ scores of VGGNet trained by 2 classes (broadleaves and grasses) reached 0.99. These findings generally indicate that dividing weeds into broadleaf and grass weeds for training the networks improved the performance of weed detection. Combined with the training samples of Sections 3.2 and 3.3, VGGNet is the most suitable neural network for detecting weeds growing in alfalfa.

Table 4. The validation and testing results using image classification and object detection neural networks to detect broadleaf weeds growing in alfalfa ^{a,b}.

Model	Weed Species	Network Type	Validation Results			Testing Results		
			Precision	Recall	F ₁ Score	Precision	Recall	F ₁ Score
GoogLeNet	<i>Artemisia vulgaris</i>	IC	0.64	0.67	0.65	0.63	0.65	0.64
	<i>Galium aparine</i>	IC	0.76	0.83	0.79	0.77	0.81	0.79
	<i>Geranium carolinianum</i>	IC	0.87	0.75	0.81	0.87	0.78	0.82
	<i>Veronica</i>	IC	0.88	0.88	0.89	0.87	0.88	0.87
VGGNet	<i>Artemisia vulgaris</i>	IC	0.95	0.90	0.92	0.95	0.89	0.92
	<i>Galium aparine</i>	IC	0.94	1.00	0.97	0.93	0.98	0.96
	<i>Geranium carolinianum</i>	IC	0.95	0.96	0.95	0.94	0.96	0.95
	<i>Veronica</i>	IC	1.00	1.00	1.00	1.00	1.00	1.00
Faster R-CNN	<i>Artemisia vulgaris</i>	OD	0.49	0.67	0.57	0.48	0.68	0.56
	<i>Galium aparine</i>	OD	0.24	0.23	0.29	0.27	0.37	0.31
	<i>Geranium carolinianum</i>	OD	0.38	0.56	0.45	0.37	0.56	0.45
	<i>Veronica</i>	OD	0.18	0.33	0.23	0.20	0.34	0.25
YOLOv3	<i>Artemisia vulgaris</i>	OD	0.90	0.78	0.84	0.88	0.72	0.81
	<i>Galium aparine</i>	OD	0.88	0.21	0.33	0.80	0.09	0.15
	<i>Geranium carolinianum</i>	OD	0.81	0.39	0.53	0.81	0.40	0.54
	<i>Veronica</i>	OD	0.54	0.29	0.36	0.79	0.23	0.35

^a Abbreviations: IC, image classification; OD, object detection. ^b The models were trained with the training dataset containing various broadleaf weeds. For the image classification networks, the training dataset contained 1000 images, while the validation dataset contained 200 images. The testing dataset contained 100 images. For the object detection networks, the training dataset contained 450 images, including a total of 647 catchweed bedstraw, 1219 Carolina geranium, 4352 mugwort, and 499 speedwell. The validation or testing dataset contained 50 images.

Table 5. Grass weed detection validation and testing results using various convolutional neural networks ^b.

Model	Weed Species	Network Type ^a	Validation Results			Testing Results		
			Precision	Recall	F ₁ Score	Precision	Recall	F ₁ Score
GoogLeNet	<i>Digitaria</i>	IC	0.47	0.55	0.50	0.42	0.55	0.48
	<i>Echinochloa crus-galli</i>	IC	0.62	0.67	0.65	0.62	0.66	0.64
	<i>Eleusine indica</i>	IC	0.64	0.47	0.54	0.51	0.31	0.39
VGGNet	<i>Digitaria</i>	IC	0.56	0.79	0.66	0.46	0.71	0.56
	<i>Echinochloa crus-galli</i>	IC	0.75	0.77	0.76	0.69	0.69	0.69
	<i>Eleusine indica</i>	IC	0.79	0.42	0.55	0.75	0.32	0.45
Faster R-CNN	<i>Digitaria</i>	OD	0.20	0.31	0.24	0.19	0.31	0.24
	<i>Echinochloa crus-galli</i>	OD	0.07	0.30	0.12	0.05	0.40	0.09
	<i>Eleusine indica</i>	OD	0.21	0.35	0.26	0.20	0.31	0.24
YOLOv3	<i>Digitaria</i>	OD	0.75	0.43	0.54	0.72	0.53	0.61
	<i>Echinochloa crus-galli</i>	OD	0.45	0.25	0.32	0.44	0.27	0.33
	<i>Eleusine indica</i>	OD	0.33	0.14	0.19	0.54	0.15	0.23

^a Abbreviations: IC, image classification; OD, object detection. ^b The training dataset contained various broadleaf weeds in order to detect various broadleaf weeds growing in alfalfa. The image classification networks of the training dataset contained 1000 images. The validation dataset contained 200 images. The testing results contained 100 images. The object detection networks of the training data set contained 560 images within which were 1008 crabgrass, 476 barnyard grass, and 497 goosegrass. The validation data set contained 60 images, and the testing results contained 60 images.

4. Discussion

In other cropping systems, Sharpe et al. [63] reported that the leaf-trained DetectNet showed a high F₁ score (0.94) for detecting Carolina geranium growing in competition with strawberry (*Fragaria × ananassa* (Weston) Duchesne ex Rozier (pro sp.) (*chiloensis × virginiana*)). Recently, to detect broadleaf weed seedlings growing in wheat (*Triticum aestivum* L.), Zhuang et al. [64] reported that object detection neural networks, including CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3, were ineffective (F₁ scores ≤ 0.68). Overall, the present study's findings suggest that the evaluated object detection neural networks may not be appropriate for detecting weeds growing in alfalfa. Further, other object detection networks such as SSD [65] and DetectNet [61] are designed to identify multiple classes per image and need to be evaluated for weed detection in future research.

The present study results showed that the image classification neural networks outperformed object detection neural networks, and GoogLeNet and VGGNet can be used to detect broadleaf and grass weeds growing in alfalfa. Similarly, Zhang et al. [48] reported that image classification neural networks, AlexNet, DenseNet, and VGGNet, outperformed the object detection neural networks, CenterNet, Faster R-CNN, TridentNet, VFNet, and YOLOv3, for the detection of broadleaf seedlings, particularly at high weed density, growing in wheat.

Image classification neural networks classify images according to the content of the images. There is usually a set of fixed categories, and the model must predict the most suitable category for the image. Object detection is more complicated than image classification because more operation and processing is involved. The difference between the lowest and highest F_1 scores of VGGNet was 0.08. However, the F_1 scores of the object detection neural networks varied greatly. The difference between the highest and the lowest F_1 scores of YOLOv3 was 0.66. These findings suggest Faster R-CNN and YOLOv3 may not be suitable for detecting weeds growing in alfalfa. For detection of mugwort, YOLOv3 outperformed Faster R-CNN, but the F_1 scores never exceeded 0.84. Overall, these findings suggest that VGGNet is the most effective neural network for detecting various broadleaf weeds while growing in alfalfa among the evaluated neural networks.

5. Conclusions

This research demonstrated that VGGNet performed well at detecting various broadleaf and grass weeds. To detect weeds growing in alfalfa, the image classification networks, GoogLeNet and VGGNet, outperformed the object detection networks, Faster R-CNN, YOLOv3, and VFNet. The VGGNet performed best in the 2 classes (broadleaves and grasses) neural network (F_1 scores ≥ 0.99) and is a suitable neural network for detecting various broadleaf and grass weeds growing in alfalfa. Overall, we conclude that the evaluated object detection networks, including Faster R-CNN, VarifocalNet (VFNet), and YOLOv3, may not be appropriate for detecting weeds growing in alfalfa. The use of VGGNet as the decision-making system for the machine vision subsystem seems to be a viable option for the precise spraying of herbicides in alfalfa.

Author Contributions: Conceptualization, methodology, J.Y. (Jie Yang) and J.Y. (Jialin Yu); software, Y.W.; validation, formal analysis, J.Y. (Jie Yang); investigation, resources, data J.Y. (Jialin Yu); writing—original draft preparation, J.Y. (Jie Yang); writing—review and editing, J.Y. (Jialin Yu); supervision, project administration, funding acquisition, J.Y. (Jialin Yu) and Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Natural Science Foundation of China (Grant No. 32072498), Key Research and Development Program of Jiangsu Province (Grant No. BE2021016) and Jiangsu Agricultural Science and Technology Innovation Fund (Grant No. CX(21)3184).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are thankful to National Natural Science Foundation of China, Key Research and Development Program of Jiangsu Province and Jiangsu Agricultural Science and Technology Innovation Fund.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hojilla-Evangelista, M.P.; Selling, G.W.; Hatfield, R.; Digman, M. Extraction, composition, and functional properties of dried alfalfa (*Medicago sativa* L.) leaf protein. *J. Sci. Food Agric.* **2017**, *97*, 882–888. [[CrossRef](#)] [[PubMed](#)]
2. Richter, N.; Siddhuraju, P.; Becker, K. Evaluation of nutritional quality of moringa (*Moringa oleifera* Lam.) leaves as an alternative protein source for Nile tilapia (*Oreochromis niloticus* L.). *Aquaculture* **2003**, *217*, 599–611. [[CrossRef](#)]

3. Salzano, A.; Neglia, G.; D'Onofrio, N.; Balestrieri, M.L.; Limone, A.; Cotticelli, A.; Marrone, R.; Anastasio, A.; D'Occhio, M.J.; Campanile, G. Green feed increases antioxidant and antineoplastic activity of buffalo milk: A globally significant livestock. *Food Chem.* **2021**, *344*, 128669. [[CrossRef](#)] [[PubMed](#)]
4. Kerr, L.A.; Johnson, B.J.; Burrows, G.E. Intoxication of cattle by *Perilla frutescens* (purple mint). *Vet. Hum. Toxicol.* **1986**, *28*, 412–416. [[PubMed](#)]
5. Cudney, D.W.; Adams, O. Improving weed control with 2,4-DB amine in seedling alfalfa (*Medicago sativa*). *Weed Technol.* **1993**, *7*, 465–470. [[CrossRef](#)]
6. Idris, K.I.; Dongola, G.M.; Elamin, S.E.; Babiker, M.M. Evaluation of clethodim for weed control in alfalfa (*Medicago sativa* L.). *Univ. Khartoum J. Agric. Sci.* **2019**, *22*, 126–135.
7. Wilson, R.G.; Burgener, P.A. Evaluation of glyphosate-tolerant and conventional alfalfa weed control systems during the first year of establishment. *Weed Technol.* **2009**, *23*, 257–263. [[CrossRef](#)]
8. Zijlstra, C.; Lund, I.; Justesen, A.F.; Nicolaisen, M.; Jensen, P.K.; Bianciotto, V.; Posta, K.; Balestrini, R.; Przetakiewicz, A.; Czembor, E. Combining novel monitoring tools and precision application technologies for integrated high-tech crop protection in the future (a discussion document). *Pest Manag. Sci.* **2011**, *67*, 616–625. [[CrossRef](#)]
9. Franco, C.; Pedersen, S.M.; Papaharalampos, H.; Orum, J.E. The value of precision for image-based decision support in weed management. *Precis. Agric.* **2017**, *18*, 366–382. [[CrossRef](#)]
10. Sabzi, S.; Abbaspour-Gilandeh, Y.; Garcia-Mateos, G. A fast and accurate expert system for weed identification in potato crops using metaheuristic algorithms. *Comput. Ind.* **2018**, *98*, 80–89. [[CrossRef](#)]
11. Yu, J.; Schumann, A.W.; Sharpe, S.M.; Li, X.; Boyd, N.S. Detection of grassy weeds in bermudagrass with deep convolutional neural networks. *Weed Sci.* **2020**, *68*, 545–552. [[CrossRef](#)]
12. Zaman, Q.U.; Esau, T.J.; Schumann, A.W.; Percival, D.C.; Chang, Y.K.; Read, S.M.; Farooque, A.A. Development of prototype automated variable rate sprayer for real-time spot-application of agrochemicals in wild blueberry fields. *Comput. Electron. Agric.* **2011**, *76*, 175–182. [[CrossRef](#)]
13. Sabzi, S.; Abbaspour-Gilandeh, Y.; Arribas, J.I. An automatic visible-range video weed detection, segmentation and classification prototype in potato field. *Heliyon* **2020**, *6*, e03685. [[CrossRef](#)] [[PubMed](#)]
14. Bakhshipour, A.; Jafari, A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* **2018**, *145*, 153–160. [[CrossRef](#)]
15. Hamuda, E.; Mc Ginley, B.; Glavin, M.; Jones, E. Automatic crop detection under field conditions using the HSV colour space and morphological operations. *Comput. Electron. Agric.* **2017**, *133*, 97–107. [[CrossRef](#)]
16. Pulido, C.; Solaque, L.; Velasco, N. Weed recognition by SVM texture feature classification in outdoor vegetable crop images. *Ing. Investig.* **2017**, *37*, 68–74. [[CrossRef](#)]
17. Farooq, A.; Hu, J.; Jia, X. Analysis of spectral bands and spatial resolutions for weed classification via deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 183–187. [[CrossRef](#)]
18. Ahmad, J.; Jan, B.; Farman, H.; Ahmad, W.; Ullah, A. Disease detection in plum using convolutional neural network under true field conditions. *Sensors* **2020**, *20*, 5569. [[CrossRef](#)]
19. Akbarzadeh, S.; Paap, A.; Ahderom, S.; Apopei, B.; Alameh, K. Plant discrimination by support vector machine classifier based on spectral reflectance. *Comput. Electron. Agric.* **2018**, *148*, 250–258. [[CrossRef](#)]
20. Sujaritha, M.; Annadurai, S.; Satheeshkumar, J.; Sharan, S.K.; Mahesh, L. Weed detecting robot in sugarcane fields using fuzzy real time classifier. *Comput. Electron. Agric.* **2017**, *134*, 160–171. [[CrossRef](#)]
21. Yu, J.; Sharpe, S.M.; Schumann, A.W.; Boyd, N.S. Deep learning for image-based weed detection in turfgrass. *Eur. J. Agron.* **2019**, *104*, 78–84. [[CrossRef](#)]
22. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)] [[PubMed](#)]
23. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
24. Ni, C.; Wang, D.; Vinson, R.; Holmes, M.; Tao, Y. Automatic inspection machine for maize kernels based on deep convolutional neural networks. *Biosyst. Eng.* **2019**, *178*, 131–144. [[CrossRef](#)]
25. Saood, A.; Hatem, I. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. *BMC Med. Imag.* **2021**, *21*, 19. [[CrossRef](#)] [[PubMed](#)]
26. Geng, D.V.; Alkhachroum, A.; Bicchi, M.A.M.; Jagid, J.R.; Cajigas, I.; Chen, Z.S. Deep learning for robust detection of interictal epileptiform discharges. *J. Neural Eng.* **2021**, *18*, 056015. [[CrossRef](#)]
27. Yao, P.; Gai, S.; Chen, Y.; Chen, W.; Da, F. A multi-code 3D measurement technique based on deep learning. *Opt. Lasers Eng.* **2021**, *143*, 106623. [[CrossRef](#)]
28. Ma, J.S.; Sheridan, R.P.; Liaw, A.; Dahl, G.E.; Svetnik, V. Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263–274. [[CrossRef](#)]
29. Ciodaro, T.; Deva, D.; De Seixas, J.; Damazio, D. Online particle detection with Neural Networks based on topological calorimetry information. In Proceedings of the 14th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT), Brunel University, Uxbridge, UK, 5–9 September 2011; IOP Publishing: Bristol, UK, 2012; Volume 368.
30. Azhari, M.; Abarda, A.; Ettaki, B.; Zerouaoui, J.; Dakkon, M. Higgs boson discovery using machine learning methods with pyspark. *Procedia Comput. Sci.* **2020**, *170*, 1141–1146. [[CrossRef](#)]
31. Shi, J.; Li, Z.; Zhu, T.; Wang, D.; Ni, C. Defect detection of industry wood veneer based on NAS and multi-channel mask R-CNN. *Sensors* **2020**, *20*, 4398. [[CrossRef](#)]

32. Zhou, H.; Zhuang, Z.; Liu, Y.; Liu, Y.; Zhang, X. Defect classification of green plums based on deep learning. *Sensors* **2020**, *20*, 6993. [CrossRef] [PubMed]
33. Helmstaedter, M.; Briggman, K.L.; Turaga, S.C.; Jain, V.; Seung, H.S.; Denk, W. Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* **2013**, *500*, 168–174. [CrossRef]
34. Liu, Y.; Sun, P.; Wergeles, N.; Shang, Y. A survey and performance evaluation of deep learning methods for small object detection. *Expert Syst. Appl.* **2021**, *172*, 114602. [CrossRef]
35. Zhao, Z.Q.; Zheng, P.; Xu, S.T.; Wu, X. Object detection with deep learning: A review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [CrossRef] [PubMed]
36. Tompson, J.; Jain, A.; LeCun, Y.; Bregler, C. Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation. In Proceedings of the 28th Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Neural Information Processing Systems (Nips): La Jolla, CA, USA, 2014; Volume 27, pp. 1799–1807.
37. Wang, Q.; Gao, J.Y.; Yuan, Y. A joint convolutional neural networks and context transfer for street scenes labeling. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 1457–1470. [CrossRef]
38. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [CrossRef]
39. Ghosal, S.; Blystone, D.; Singh, A.K.; Ganapathysubramanian, B.; Singh, A.; Sarkar, S. An explainable deep machine vision framework for plant stress phenotyping. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4613–4618. [CrossRef]
40. Singh, A.K.; Ganapathysubramanian, B.; Sarkar, S.; Singh, A. Deep learning for plant stress phenotyping: Trends and future perspectives. *Trends Plant Sci.* **2018**, *23*, 883–898. [CrossRef]
41. Wang, A.C.; Zhang, W.; Wei, X.H. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* **2019**, *158*, 226–240. [CrossRef]
42. Sharpe, S.M.; Schumann, A.W.; Yu, J.; Boyd, N.S. Vegetation detection and discrimination within vegetable plasticulture row-middles using a convolutional neural network. *Precis. Agric.* **2020**, *21*, 264–277. [CrossRef]
43. Hennessy, P.J.; Esau, T.J.; Farooque, A.A.; Schumann, A.W.; Zaman, Q.U.; Corscadden, K.W. Hair fescue and sheep sorrel identification using deep learning in wild blueberry production. *Remote Sens.* **2021**, *13*, 943. [CrossRef]
44. Hussain, N.; Farooque, A.A.; Schumann, A.W.; Abbas, F.; Acharya, B.; McKenzie-Gopsill, A.; Barrett, R.; Afzaal, H.; Zaman, Q.U.; Cheema, M.J.M. Application of deep learning to detect lamb's quarters (*Chenopodium album* L.) in potato fields of Atlantic Canada. *Comput. Electron. Agric.* **2021**, *182*, 106040. [CrossRef]
45. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; IEEE: New York, NY, USA, 2015; pp. 1–9.
46. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 2nd International Conference on Learning Representations, Banff, AB, Canada, 14–16 April 2014.
47. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Proceedings of the 29th Annual Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Neural Information Processing Systems (Nips): La Jolla, CA, USA, 2015; Volume 28, pp. 91–99.
48. Zhang, H.; Wang, Y.; Dayoub, F.; Sunderhauf, N. Varifocalnet: An IoU-aware dense object detector. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Electrical Network, Nashville, TN, USA, 20–25 June 2021; IEEE Computer Society: Los Alamitos, CA, USA; pp. 8514–8523.
49. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
50. Zeiler, M. ADADELTA: An Adaptive Learning Rate Method. *arXiv* **2012**, arXiv:1212.5701.
51. Hochreiter, S.; Bengio, Y.; Frasconi, P.; Schmidhuber, J. Gradient flow in recurrent nets: The difficulty of learning long-term dependencies. In *A Field Guide to Dynamical Recurrent Neural Networks*; Kolen, J.F., Kremer, S.C., Eds.; IEEE Press: Piscataway, NJ, USA, 2001; ISBN 978-0-7803-5369-5.
52. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef] [PubMed]
53. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.F. ImageNet: A large-scale hierarchical image database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, USA, 20–25 June 2009; IEEE: New York, NY, USA, 2009; Volume 1–4, pp. 248–255.
54. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
55. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
56. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the ACM Conference on Multimedia, Orlando, FL, USA, 3–7 November 2014; ACM: New York, NY, USA, 2014; pp. 675–678.
57. Chen, K.; Wang, J.; Pang, J.; Cao, Y.; Xiong, Y.; Li, X.; Sun, S.; Feng, W.; Liu, Z.; Xu, J. MMDetection: Open MMLab Detection Toolbox and Benchmark. *arXiv* **2019**, arXiv:1906.07155.
58. Redmon, J. Darknet: Open Source Neural Networks in C (2013–2016). Available online: <https://pjreddie.com/darknet/> (accessed on 10 September 2018).

59. Hoiem, D.; Chodpathumwan, Y.; Dai, Q. Diagnosing error in object detectors. In Proceedings of the 12th European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7574, pp. 340–353.
60. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Processing Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
61. Tao, A.; Barker, J.; Sarathy, S. Detectnet: Deep Neural Network for Object Detection in DIGITS. Available online: <https://devblogs.nvidia.com/detectnet-deep-neural-network-object-detection-digits> (accessed on 11 May 2018).
62. Milioto, A.; Lottes, P.; Stachniss, C. Real-time blob-wise sugar beets vs. weeds classification for monitoring fields using convolutional neural networks. In Proceedings of the International Conference on Unmanned Aerial Vehicles in Geomatics, Bonn, Germany, 4–7 September 2017; Copernicus Gesellschaft MBH: Gottingen, Germany, 2017; Volume 4, pp. 41–48.
63. Sharpe, S.M.; Schumann, A.W.; Boyd, N.S. Detection of carolina geranium (*Geranium carolinianum*) growing in competition with strawberry using convolutional neural networks. *Weed Sci.* **2019**, *67*, 239–245. [[CrossRef](#)]
64. Zhuang, J.; Li, X.; Bagavathiannan, M.; Jin, X.; Yang, J.; Meng, W.; Li, T.; Li, L.; Wang, Y.; Chen, Y. Evaluation of different deep convolutional neural networks for detection of broadleaf weed seedlings in wheat. *Pest Manag. Sci.* **2021**, *78*, 521–529. [[CrossRef](#)]
65. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 21–37.