



A Review of Deep Learning in Multiscale Agricultural Sensing

Dashuai Wang ¹, Wujing Cao ¹, Fan Zhang ², Zhuolin Li ³, Sheng Xu ¹, and Xinyu Wu ^{1,*}

- Guangdong Provincial Key Laboratory of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518005, China; ds.wang1@siat.ac.cn (D.W.); wj.cao@siat.ac.cn (W.C.); sheng.xu@siat.ac.cn (S.X.)
- ² College of Engineering, China Agricultural University, Beijing 100083, China; BS20193070612@cau.edu.cn
- ³ School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China; zl.li@siat.ac.cn
- * Correspondence: xy.wu@siat.ac.cn

Abstract: Population growth, climate change, and the worldwide COVID-19 pandemic are imposing increasing pressure on global agricultural production. The challenge of increasing crop yield while ensuring sustainable development of environmentally friendly agriculture is a common issue throughout the world. Autonomous systems, sensing technologies, and artificial intelligence offer great opportunities to tackle this issue. In precision agriculture (PA), non-destructive and non-invasive remote and proximal sensing methods have been widely used to observe crops in visible and invisible spectra. Nowadays, the integration of high-performance imagery sensors (e.g., RGB, multispectral, hyperspectral, thermal, and SAR) and unmanned mobile platforms (e.g., satellites, UAVs, and terrestrial agricultural robots) are yielding a huge number of high-resolution farmland images, in which rich crop information is compressed. However, this has been accompanied by challenges, i.e., ways to swiftly and efficiently making full use of these images, and then, to perform fine crop management based on information-supported decision making. In the past few years, deep learning (DL) has shown great potential to reshape many industries because of its powerful capabilities of feature learning from massive datasets, and the agriculture industry is no exception. More and more agricultural scientists are paying attention to applications of deep learning in image-based farmland observations, such as land mapping, crop classification, biotic/abiotic stress monitoring, and yield prediction. To provide an update on these studies, we conducted a comprehensive investigation with a special emphasis on deep learning in multiscale agricultural remote and proximal sensing. Specifically, the applications of convolutional neural network-based supervised learning (CNN-SL), transfer learning (TL), and few-shot learning (FSL) in crop sensing at land, field, canopy, and leaf scales are the focus of this review. We hope that this work can act as a reference for the global agricultural community regarding DL in PA and can inspire deeper and broader research to promote the evolution of modern agriculture.

Keywords: precision agriculture; deep learning; convolutional neural networks; transfer learning; few-shot learning; remote sensing; proximal sensing

1. Introduction

The Food and Agriculture Organization (FAO) of the United Nations has projected that the world's growing population will reach nine billion people and global food production will need to almost double between 2005 and 2050 [1]. However, climate change is contributing to an increasing frequency of extreme weather conditions that are disastrous for worldwide agriculture [2,3]. Unfortunately, the arrival of the COVID-19 pandemic at the end of 2019 has created more uncertainties and risks to global food security [4]. The use of technical means to deal with these ever-increasing challenges has become an important research direction in the agricultural community.



Citation: Wang, D.; Cao, W.; Zhang, F.; Li, Z.; Xu, S.; Wu, X. A Review of Deep Learning in Multiscale Agricultural Sensing. *Remote Sens.* 2022, *14*, 559. https://doi.org/ 10.3390/rs14030559

Academic Editors: Xiuyuan Zhang, Shihong Du and Gong Cheng

Received: 15 December 2021 Accepted: 23 January 2022 Published: 25 January 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/).

Over the past few decades, significant achievements in sensing technologies, wireless communication, autonomous systems, and artificial intelligence have been made by research efforts from all over the world [5,6]. This is good news for agriculture because these technological developments are significant to the promotion of sustainable agricultural development. Currently, agricultural scientists are dedicated to evolving traditional agriculture into data-driven precision agriculture (PA), with the main goal being to apply the right treatment in the right place at the right time while taking the spatiotemporal variabilities of farmland into account [7]. Farmland information acquisition and processing are two indispensable prerequisites for PA. In terms of farmland information acquisition, high-performance sensors, such as RGB, multispectral, hyperspectral, thermal, and SAR cameras have been used to observe physical and physiological changes in plants in both visible and invisible light [8]. In recent years, the agriculture sector has witnessed a considerable increase in civilian satellites, unmanned aerial vehicles (UAVs), and autonomous field robots [9–11]. Aerial or terrestrial platforms in conjunction with high-performance sensors are yielding tremendous amounts of farmland images with various temporal, spatial, and spectral resolutions. More importantly, these images contain a lot of meaningful information that is expected to help farmers to schedule sowing, track the growth status of their crops, monitor pests and diseases, control weeds, and predict yield [12]. However, rapid and accurate extraction of the most valuable information from large amounts of imagery data that have been collected under various conditions is becoming an increasingly rigorous challenge that is being faced by agricultural scientists as well as rural farming communities [13].

Along with the flourish of image processing, computer vision, and artificial intelligence, there has been significant progress made toward learning-based feature extraction strategies [14–16]. The most popular techniques used for analyzing images include machine learning (ML) (K-means and support vector machine (SVM), artificial neural networks (ANN), amongst others), linear polarizations, wavelet-based filtering, and regression analysis. Moreover, deep learning (DL) [17], which is used to extract hierarchical features from images, has unlocked fresh prospects for interpretating enormous amounts of data and has percolated into data analytics in agriculture [18–20]. Within the context of PA, agricultural scientists have shown great interests in transferring the successful advancements of DL that have been made in non-agricultural applications to agricultural applications [21–23]. This is motivated by the intuitive judgment that a deeper understanding and consecutive tracking of farmland information is conducive to more scientific decision making and management. For instance, in plant stress monitoring, DL-enhanced image classification, detection, segmentation, and tracking is essential to understand "what" and "where" pathogen-insect-plant interactions are, "how" plant stresses occur, distribute, and develop, and "when" human intervention is needed.

Several studies have been conducted that have surveyed DL applications in PA. One of the most cited works is a review study conducted by Kamilaris, A. and Prenafeta-Boldú, F.X. (2018) [18]. This review covered 40 research efforts that employed DL and discussed a wide range of topics, including architectures, frameworks, algorithms, datasets, metrics, applications, challenges, limitations, and perspectives. Although this work was relatively broad, it demonstrated the required elements and outlined the baseline framework of agricultural DL. Singh, A. and colleagues outlined some machine learning (ML) applications and their deep learning subtypes for image-based high-throughput plant stress phenotyping [12,21,22]. They compared the advantages and disadvantages of ML and DL in high-throughput stress phenotyping and introduced some concepts that simplified transition to practice, such as transfer learning (TL) and active learning (AL). Additionally, they also surveyed relevant works at multiple scales, such as leaf scale stress phenotyping, plant canopy and small-plot stress phenotyping, and field scale foliar stress phenotyping. Specifically focused on DL-based weed stress detection, Hasan, A. S. M. M. et al. performed a comprehensive survey of 70 works related to automatic detection and classification of weeds in crops and elaborated existing deep learning techniques from five aspects, namely data acquisition, dataset preparation, detection approaches, learning methods, and DL architectures [24]. This work concluded that supervised learning based on labeled datasets was commonly used and that relatively high accuracy was achieved with limited experiment setup. Nowadays, data have become the primary prerequisite for various DL applications in PA. In 2020, Lu, Y. and Yong, S. investigated publicly available image datasets for computer vision tasks in PA [25]. A total of 34 image datasets were identified and classified into three categories: 15 datasets for weed control, 10 datasets for fruit detection, and 9 datasets for other applications. For each dataset, key information was summarized such as number of images, image acquisition method, image type, image size, image resolution, image processing method, and application purpose. However, it is worth mentioning that the quantity, diversity, complexity, and quality of currently available datasets are not sufficient to support the widespread promotion of deep learning in agriculture. Agricultural scientists have proposed some other strategies, such as transfer learning (TL) [26] and few-shot learning (FSL) [27], to reduce the dependence of deep learning models on datasets. For example, TL has been successfully used for weed and disease identification [28–30], and FSL has been shown to be helpful for plant disease classification and recognition [31–33].

Over the past few years, deep learning has been widely deployed in various agricultural applications for identification, classification, detection, quantification, and prediction. Specifically, as shown in Figure 1, DL can be used for disease and pest identification and quantification at the leaf scale; for plant-weed classification and identification at the canopy scale; and for abiotic stress assessment, plant growth and nutrient status monitoring, and yield prediction at the field scale; while at the land scale, DL can be implemented for land cover mapping and crop discrimination. There is no doubt that previous studies have made significant contributions to DL in PA by discussing relevant applications, key achievements, potential challenges, and future perspectives. Nevertheless, none of the studies have focused on DL applications in multiscale agricultural sensing throughout the entire crop growth cycle.



Figure 1. Deep learning in multiscale agricultural sensing.

Motivated by the rapid development of precision agriculture and deep learning, in this work, we have conducted a comprehensive review on the applications of DL for multiscale agricultural remote/proximal sensing. The term "deep learning", as it is used in this work, is particularly limited to agricultural image processing techniques that are based on deep neural networks, such as classification, recognition, and segmentation. The term "precision agriculture" specifically refers to the acquisition, analysis, and decision making related to the farmland where crops are cultivated. Plant stresses, such as weeds, diseases, and pests, specifically target those stresses that can be detected by the electromagnetic spectrum. The most significant contributions of this work are highlighted as follows:

- (i). According to learning strategies, we broadly classified the DL frameworks that are commonly used in PA into three categories, i.e., CNN-based supervised learning (CNN-SL), transfer learning (TL), and few-shot learning (FSL).
- (ii). We conducted a comprehensive review of typical studies on the application of CNN-SL, TL, and FSL, at the leaf, canopy, field, and land scales of agricultural sensing.
- (iii). We elaborated the limitations that impede the application of DL in PA, discussed the current challenges, and we proposed our perspectives for future research.

In the most popular scientific databases, such as IEEE Xplore, ScienceDirect, Web of Science, and Google Scholar, we used keywords such as "precision agriculture", "deep learning", and "remote sensing" to acquire relevant journal articles and conference papers. Then, these materials were carefully studied and classified to form this review. The remainder of the paper is organized as follows: in Sections 2–4, we focus on the use of CNN-SL, TL, and FSL in PA, respectively, and specifically, in Section 2, CNN-SL-based deep learning in PA at the leaf, canopy, field, and land scales is discussed in detail; in Section 5, we discuss the challenges and opportunities for DL in PA; and finally, in Section 6, we summarize the findings of the previous sections and deliver some final remarks.

2. CNN-Based Supervised Learning (CNN-SL) in PA

The state-of-the-art convolutional neural network-based supervised learning (CNN-SL) is the most popular framework among deep learning models applied in precision agriculture, as it can quickly and accurately extract the most effective characterization information from various features (color, shape, texture, size, spectral, venation, etc.) at different levels without human expertise [34,35]. One of the most significant features of CNN-SL is that it relies on sufficient datasets with annotated labels. Intensive dataset collection and manual labeling are common practices for DL-based applications in PA. With contributions from the agricultural research community, many purpose-specific datasets have been built and shared publicly. Meanwhile, some state-of-the-art CNN models have been introduced in PA for farmland observation and management at multiple scales.

2.1. Leaf Scale: Disease and Pest Identification and Quantification

Early diagnosis and the instant control of plant diseases and pests is extremely important for reducing yield loss. However, speedy identification of biotic stresses and accurate quantification of threatening severity is a significant challenge, given their individual differences and spatial-temporal diversity [36,37]. The use of CNN-SL for developing low cost, general purpose, automatic, and accurate plant disease and pest classification and quantification algorithms has been explored extensively in recent years [38–40].

PlantVillage [41], an openly accessible dataset released in 2015, has contributed to many DL-based agricultural studies. It contains 54,309 images of leaves of healthy and diseased plants. For each plant, a leaf was picked and then placed on a paper sheet that offered a clear background. Subsequently, from four to seven images were collected at a standard point by a digital camera (Sony DSC-Rx100/13, 20.2 megapixels, automatic mode) under natural light. All the images were labeled by plant pathology experts. Aiming to develop a smartphone-assisted plant disease diagnosis system, Mohanty, S.P. et al. (2016) trained two well-known deep convolutional neural networks, AlexNet and GoogLeNet, to identify 14 crop species and 26 diseases in PlantVillage; GoogLeNet achieved the top

accuracy of 99.35%. However, when tested in two other validation datasets (IPM and Bing), the overall accuracy reduced substantially, to just over 31%, which means the model's generalization ability was not as good as expected [42]. Ferentinos, K.P. (2018) trained and assessed five basic CNN architectures to identify plant diseases by using an extended version of PlantVillage, which contained 87,848 images captured in laboratory conditions and also in cultivation fields [43]. Amongst the other CNN architectures, the VGGNet achieved the highest success rate of 99.53% in the testing set. However, due to the limited classes of plants and diseases in this dataset, it was still quite far from being a generic approach that could be deployed in real conditions. Inspired by the above two works, Lee, S.H. et al. (2020) compared the performance of different learning strategies based on two pretraining tasks and three representative network architectures (VGG16, InceptionV3, and GoogLeNet) and proved, once again, that the VGG16 model generalized better to unseen data [44]. They discovered a new phenomenon, i.e., networks trained using the crop-disease pair may not always focus on disease regions, but instead, focus on cropspecific characteristics so that the generalization of the model may be limited. Therefore, instead of using the crop-disease pair, they presented a more intuitive method that focused on diseases independently, and indicated that it was more effective, particularly, when distinguishing diseases from crops that were not included in the training set.

Picon, A. et al. (2019) created a new dataset for massive multicrop plant disease classification [45,46]. This dataset comprised a total number of 121,955 images captured by cell phone in real field conditions during consecutive growing seasons, and covered almost equally distributed disease stages (initial stages, early stages, medium stages, and advanced stages each account for a quarter) of 17 fungal diseases on 5 crops. The images were all taken with an appropriate focus, in flash mode off configuration of the camera. After being segmented by a fully convolutional DenseNet network, each image was cropped to the region surrounding the main leaf. They also proposed a novel crop-conditional classification method that seamlessly incorporated additional contextual metadata (such as crop identification, weather conditions, and geographical location) onto a ResNet50-based convolutional neural network. Comparative experiments demonstrated that this crop conditional CNN architecture obtained richer and more robustly shared visual features and yielded an average balance accuracy of 98%.

Although convolutional neural networks are usually used for object classification and detection in 2D RGB images, their principles can also be extended to feature extraction in 3D images that contain high-resolution spectral and spatial information in invisible wavelengths [47,48]. To identify charcoal rot disease in soybean stems, Nagasubramanian, K. et al. (2019) established a 3D dataset containing 111 hyperspectral images of size $500 \times 1600 \times 240$ (height \times length \times spectral frequency) and deployed a supervised 3D deep convolutional neural network to learn the spatial-temporal features [49]. Images of healthy and infected soybean stems at 3, 6, 9, 12, and 15 days after infection were captured by a Pika XC hyperspectral line imaging scanner set in a laboratory. Each image had a 2.5 nm spectral resolution in 240 wavebands over a spectral range of 400–1000 nm. Data augmentation by extracting data patches was used to address the problem of insufficient datasets. Ultimately, the training, validation, and test dataset consisted of 1090, 194, and 539 images ($64 \times 64 \times 240$), respectively. The results demonstrated that their model achieved a classification accuracy of 95.73%. However, authors also made it clear that enormous computing power was needed to interpret high dimensional 3D dataset.

Obviously, if a disease can be identified but the severity cannot be quantitatively analyzed, then, the requirements of precision agriculture are far from being achieved. Focused on the quantification of leaf spot disease severity in sugar beet, Ozguven, M. M. et al. (2019) updated a Faster R-CNN model by changing the parameters of CNN architecture and developed a new scale that was able to classify the disease severity as healthy, mild, severe, or mild and severe mixed [50]. The authors claimed that this work achieved an accuracy of 95.48% for disease detection and classification in 155 sugar beet leaf images. To assess the disease severity of powdery mildew on cucumber leaves, a

DL-based pixel-wise semantic segmentation model was implemented by Lin, K. et al. (2019) [51]. It is worth pointing out that the study used only 50 images captured by the laboratory phenotyping platform. Therefore, although the model achieved an average pixel accuracy of 96.08%, its performance in real field remains to be verified. For the automatic quantification of Northern Leaf Blight (NLB) disease in field maize, Garg, K. et al. (2021) defined the Severity Index (SI) to represent the disease severity ratio and proposed an end-to-end deep learning framework, called Cascaded MRCNN [52]. The dataset contained 7669 field maize images that had been taken, during the summer of 2017, by mounting a camera on a DJI Matrice 600 UAV. The UAV was flown at a height of 6 m and a forward speed of 1 m/s, and images were captured with nadir view every 2 s. The diseased regions were manually labeled with the help of a VGG Image Annotator. The experimental results on a field maize dataset with NLB disease showed a correlation between the predicted and ground truth SI of 73%.

Additionally, CNN-SL represents a promising solution for crop pest detection. Many researchers are committed to truly applying deep learning detection methods to actual farmland environments. However, field pest detection and recognition has encountered significant challenges considering the complexity of the farmland environment, leaf overlap occlusion, tiny but diverse pest characteristics, etc. Li, Y. et al. (2020) put forward a DLbased crop pest recognition framework that was able to categorize 10 common crop pests in natural scenes [53]; they collected 5629 RGB images by downloading images from search engines and performing field shooting. By applying offline data augmentation methods, such as rotation, mirroring, noise addition, and zooming, the number of images in the original dataset was increased to 14,475. In this study, the performances of five different CNN models (VGG-16, VGG-19, ResNet50, ResNet152, and GoogLeNet) were investigated. The test results indicated that GoogLeNet performed better than other models in terms of accuracy, model complexity, and robustness. In Cheng, X. et al. (2017), a deep residual learning agricultural pest identification approach that achieved a classification accuracy of 98.67% for 10 crop pest classes under a complex farmland background was proposed [54]. Towards the high-accuracy detection and recognition of multiple species of pest within the field at a large scale, Wang, F. et al. (2020) introduced a task-specific in-field pest dataset named In-Field Pest in Food Crop (IPEC) and a novel two-stage DL-based pest detection approach called DeepPest [55]. They collected 17,192 pest images in IPEC, in 2017 and 2018, by Chinese research with a Sony CX-10 CCD camera whose parameters were set to 4 mm focal length with an aperture of f/3.3. For the process of labeling data, first, human experts were employed to label a few images for training a multiscale context-aware information representation method, and then the trained model was used to automatically label more pest images. DeepPest has two stages. The first stage takes multiscale contextual information as prior knowledge and constructs a context-aware attention network for crop classification. The second stage uses a multiprojection pest detection model to generate the super-resolved feature. Experimental results on their own dataset have shown that DeepPest outperformed state-of-the-art object detection methods in pest detection within real field situations.

2.2. Canopy Scale: Plant and Weed Classification, Identification, and Positioning

The automatic and accurate awareness of specific species, growth status, distribution density, and spatial location of field plants (including crops and weeds) are the preconditions for agricultural robots to perform successive crop monitoring and specific management [56–59]. CNN-SL can be used to classify and locate individuals within a single plant species, and it can also be used to discriminate crops and weeds between species [60].

Dyrmann, M. et al. (2016) constructed a deep convolutional neural network to perform a fine-grained separation of seedlings in their respective species [61]. A total of 10,413 images (128×128) containing 22 weed and crop species at early growth stages were collected by laboratory camera and handheld mobile phones in fields to train the network. All images were photographed vertically towards the ground. Data augmentation was also applied in this work. By rotating and mirroring the original training data, 50,864 training samples were obtained. The test results indicated that the average classification accuracy of this network was 86.2%. Although the accuracy and robustness of this network was not outstanding from the current point of view, this work verified the feasibility of neural network applications in the agricultural field and provided a reference for subsequent work. Lottes, P. and his colleagues made significant contributions to the literature on the visionbased classification of plants and weeds [62]. Recently, they also presented great interest in applying deep learning in precision agriculture. For instance, in studies by Lottes, P. et al. (2020 and 2018), a fully convolutional network (FCN) was proposed to estimate plant or weed stem position and to perform pixel-wise crop-weed-soil semantic segmentation. In this study, for data acquisition and deployment, a BoniRob agricultural robot with a fourchannel JAI AD-130 GE camera mounted under a shaded area and pointed downwards on the field approximately 70–85 cm. All the images (1296 \times 966) were collected in artificial light setup inside the shaded area from different fields located in different cities in different countries with a frequency of 1 Hz, while the robot was moving over the field with a speed of approximately 0.3 m/s. Experiments proved that this approach could be effectively generalized to unknown field scenes despite dynamic changes in environmental conditions. Additionally, this work also increased stem detection accuracy and improved semantic segmentation performance [63,64]. In an attempt to classify weed and crop species through the use of convolutional neural networks, several specific network architectures have been proposed. For example, Chavan, T. R. and Nandedkar, A. V. (2018) developed AgroAVNET, a convolutional neural network derived from AlexNet and VGGNet, to identify different plant species of crops and weeds [65]. A publicly available dataset that contained a total of 407 images of 12 species of plants was used to evaluate the proposed method. The authors claimed that AgroAVNET outperformed AlexNet and VGGNET both in terms of average accuracy and in terms of training efficiency. Moreover, to improve weed and crop recognition accuracy with small amounts of labeled data in challenging field environments, Jiang, H. et al. (2020) constructed a semisupervised graph convolutional network (GCN) called GCN-ResNet-101 [66]. Four datasets, namely corn, lettuce, radish, and mixed weed were used, of which corn and lettuce were established by the authors. A total of 6000 weed and corn images (800×600) in corn dataset were taken, in 2016, using a Canon PowerShot SX600 HS camera vertically facing the ground in the natural field; 500 lettuce seedling images and 300 weed images were acquired from a 30 cm height. This novel approach outperformed conventional CNN models (AlexNet, VGG16, and ResNet-101) by achieving 97.80%, 99.37%, 98.93%, and 96.51% recognition accuracies on four different datasets, respectively.

Weeds are among the major factors that could harm crop yield [67]. Weed identification is one of the most promising domains where deep learning can significantly outperform conventional manual inspection approaches and handcrafted feature-based machining learning methods. From 2017 to 2018, Olsen, A. et al. built DeepWeeds, a dataset consisting of 17,509 annotated images of eight weed species from the Australian rangelands. To accelerate the image collection process, they developed a datalogger system, named WeedLogger. At a working distance of 1 m, this optical system provided a 4 px per mm resolution. Thereafter, a baseline for weed classification on the dataset using two deep learning models (Inception-v3 and ResNet-50), which achieved average classification accuracies of 95.1% and 95.7%, respectively, was also presented [68]. With the DeepWeeds dataset, Hu, K. et al. (2020) proposed a novel graph-based deep learning architecture, namely Graph Weeds Net (GWN), to classify multiple types of weeds from RGB images. The weighted average precision of the proposed GWN was improved from 95.7% to 97.4% for ResNet-50 [69]. Another dataset was created by dos Santos Ferreira, A. et al. (2017), i.e., the Grass-Broadleaf, which consisted of 400 images (4000×3000) of soil, soybean, and weeds [70]. A UAV (DJI Phantom 3 Professional) companioned with a Sony EXMOR 1/2.3" pointing in the nadir direction was deployed to capture images, at an average altitude of 4 m above ground level, between eight and ten in the morning,

at least once a week, during December 2015 and March 2016. Using the super pixel segmentation algorithm SLIC, 15,336 images (3249 of soil, 7376 of soybean, 3520 of grass, and 1191 of broadleaf weeds) were segmented from the original dataset through the Pynovisão software. Next, they compared the capability of convolutional neural networks to detect weeds in soybean crops with three traditional classifiers. The test results demonstrated that their work achieved accuracy above 98% in the detection of broadleaf and grass weeds. The abovementioned architectures can be classified into supervised deep neural networks, as they rely on time-consuming manual data labeling. In addition, dos Santos Ferreira, A. et al. (2019) were interested in unsupervised deep learning applications for weed discrimination [71]. With the above two datasets (Grass-Broadleaf and DeepWeeds), they evaluated the performance of two unsupervised deep clustering algorithms and proposed an approach where semi-automatic data labeling was used to generated annotations. As a result, their method achieved 97% accuracy in grass and broadleaf discrimination, while reducing the number of manual annotations by 100 times. However, semi-supervised learning and unsupervised learning in agriculture are not covered in our review.

2.3. Field Scale: Abiotic Stress Assessment, Growth Monitoring, and Yield Prediction

Apart from crop/weed classification and disease/pest identification, CNN-based supervised learning has also reshaped PA in other ways, such as crop abiotic stress assessment, plant growth and nutrient deficit monitoring, and yield prediction.

The occurrence of several abiotic stresses (water, salinity, or heat) is lethal to crops [72]. Some symptoms of plant response under abiotic stress can be presented in images both in visible and invisible wavebands. Consequently, the combination of remote/proximal sensing technologies and deep learning can play an important role in detecting the occurrence and in assessing the severity of abiotic stress [73]. For instance, Chandel, N. S. et al. (2020) trained three different CNN-SL models, namely AlexNet, GoogLeNet, and Inception V3, without modifications to classify the images between water stressed and unstressed [74]. The dataset used in this work was composed of 1200 RGB images (2500×2611) of each crop (maize, okra, and soybean). All these images were captured at different stages of growth during the seasons of 2018–2020 under natural illumination conditions at 12:00 using a Canon SX170 IS digital camera (maximum resolution of 16 megapixel, 1/2.3-inch CCD sensor, Canon DIGIC 4 image processor with 28 mm wide-angle lens having 1–1/3200 s shutter speed, 1600 maximum ISO sensitivity, and f/3.5-f/5.9 aperture). The camera was stabilized by a tripod at different distances, keeping the angle between the camera lens and the crop axis at $45 \pm 5^{\circ}$. The authors claimed that GoogLeNet outperformed the other two models with accuracies of 98.3, 97.5, and 94.1% for identifying stress or non-stress on the three crops, respectively. Moreover, to obtain high-throughput plant salt-stress phenotyping, Feng, X. et al. (2020) introduced deep learning and hyperspectral imaging into okra salinity stress assessment [75]. The 3D hypercubes data ($672 \times n \times 512$) was collected in a dark room by a linear hyperspectral scanner which was mounted 24 cm above the okra canopy. Plant phenotypes of 13 okra genotypes after 2 and 7 days of salt treatment were investigated from March to May 2018 over a period of 67 days. Specifically, a deep learning model was used to segment both whole plant and individual leaves from the background. This approach achieved an IoU of 0.94 for plant segmentation and a symmetric best dice score of 85.4 for leaf segmentation.

Plant growth stages and nutrient status monitoring play an important role in the management schedule and in yield prediction [76,77]. Currently, the most common means still rely on expert visual inspection or laboratory-based chemical analysis, which is labor-intensive and time-consuming, but also impractical within the field in large-scale conditions. Machine learning-based image processing techniques offer a promising opportunity for continuous monitoring of plant growth stages and nutrient status at the field level during the whole life cycle of a crop [78]. For example, Rasti, S. et al. (2020) investigated the performance of three different machine learning approaches (CNN model, VGG19-based transfer learning, and SVM) for estimating wheat and barley growth stages with their

customized dataset [79]. Over two consecutive years (2018 and 2019), the authors used a cell phone and a DJI Osmo+ camera to record high-quality videos at a height of 2 m in two postures, vertically downward looking at the field and at 45° declination from the horizon; 138,000 images of 12 wheat growth stages and 11 barley growth stages were extracted from video recordings with a minimum of 120 ms. The experimental results showed that CNN-based models outperformed conventional SVM classifiers, and further, VGG19-based transfer learning achieved the highest accuracy for two crops and reduced training time as well. Since convolutional neural networks are expert in spatial feature extraction, while long short-term memory networks are suitable for bridging temporal dependencies between sequential features, Abdalla, A. et al. (2020) developed a hybrid framework by concatenating a convolutional neural network with long short-term memory (CNN-LSTM) for oilseed rape nutrient diagnosis [80]. In terms of their dataset, a tripodmounted Canon camera (PowerShot SX720 HS) was used to collect RGB images at 1.2 m above the ground surface facing downward under sunny and partly cloudy conditions in the morning hours (8:30–12:00) local time. During 2017/2018, 700 images were collected at each of three growth stages, while during 2018/2019, 1500 images were collected at each of four growth stages. Therefore, this dataset was comprised of 8100 sequential infield images of the crops at different growth stages. As compared with high-dimensional handcrafted features, the feature extraction performance of five well-known CNN-SL models, VGG16, AlexNet, ResNet18, Inception3, and ResNet101 were investigated. All these networks were pretrained on the ImageNet. Then the Pretrained CNNs were finetuned on this dataset to learn features that are fed to the LSTM. This study suggested that Inceptionv3-LSTM secured the highest overall classification accuracy of 95% when using a cross-dataset validation and maintained a relatively great generalization capability to differentiate the symptoms of different nutrient levels. This study provides a baseline to launch vision-guided mobile platforms for real-time plant monitoring for improving crop production. Additionally, DL-based remote sensing can provide strong support for the accurate and continuous assessment of plant lodging. In a study by Zhang, D. et al. (2020), a new method integrating transfer learning and the DeepLab3+ network was proposed to estimate and further predict wheat lodging at different growth stages with RGB and multispectral images [81].

Maximum crop yield at minimum cost with a healthy ecosystem is one of the main goals of PA [82]. Deep learning, including CNN and RNN, is a powerful tool that can be used to make a timely prediction of crop yield from RGB or spectral images [83,84]. Aiming to develop a mobile device friendly automated system for cotton yield prediction, Tedesco-Oliveira, D. et al. (2020) investigated the performance of three detection algorithms, namely Faster R-CNN, SSD, and SSDLITE [85]. The comparative results showed that SSD achieved the lowest mean percentage errors of 8.84%. Instead of comparing the performances of various well-known deep learning models, Nevavuori, P. et al. (2019) built a CNN-based model for wheat and barley yield prediction using RGB images and NDVI images and focused on investigating the influence of network architectures and parameters on the final prediction accuracy [86]. Multispectral data were collected by a UAV-borne (Airinov Solo 3DR) multispectral camera (Parrot's SEQUIOA) during the growing season of 2017. The resolution of the UAV data was 0.3125 m. A Pix4DMapper was deployed to stitch all geolocated images in individual spectral bands acquired on a single day to form complete orthomosaic images of RGB and NDVI. The results indicated that the proposed CNN architecture performed better with RGB data than the NDVI data, and with the optimal configuration, this model could predict within-field yield with a mean absolute percentage error of 8.8% based only on RGB images in the early growth stages. Additionally, Chu, Z. et al. (2020) revealed their end-to-end model which fused two back-propagation neural networks (BPNNs) and an independently recurrent neural network (IndRNN), called the BBI model, to predict summer and winter rice yields with the fusion of deep spatial and temporal features from the area and time series meteorology data [87]. The experimental

results demonstrated that the BBI model converged quickly and achieved the best prediction performance for both summer and winter rice yields.

2.4. Land Scale: Land Cover Mapping

With the advent of spaceborne remote sensing technologies, satellite imagery has been widely used in agriculture. Thanks to public and private sector investments, the spatial resolution, temporal frequency, and spectral availability of satellite imagery have increased substantially [9]. However, fully understanding sense of these valuable images and turning them into practice guidance for agricultural production is a challenging topic within the agricultural community. Some researchers have shown great interest in introducing state-of-the-art deep learning into satellite-based farmland mapping, crop classification, yield prediction, and so on [88–90].

Papadomanolaki, M. et al. evaluated the performance of certain state-of-the-art deep learning models and their ability to classify multispectral remote sensing data. As compared with deep belief networks, the autoencoders and semi-supervised frameworks of the AlexNet and VGG models delivered a classification accuracy of over 99% [91]. With the rapid advances in network architectures, more deep layers and useful tricks have been added to benchmark frameworks to eliminate the effects of overfitting and to promote classification accuracy. Sagan, V. et al. conducted a study that directly utilized raw multispectral and temporal satellite (WorldView-3 and PlanetScope) imagery for soybean and corn yield prediction [92]. In this work, satellite-borne and UAV-borne remote sensing imagery was collected in 2017. Specifically, 4 sets of cloud-free WorldView-3 imagery and 25 sets of cloud-free PlanetScope imagery were downloaded. Additionally, a lightweight UAV (DJI Mavic Pro) with a flight height of 30 m and front/side overlap of 85%, was deployed to collect visual imagery. High-resolution UAV imagery can also be used to accurately geolocate each harvest plot on satellite imagery. An end-to-end deep learning approach using 2D and 3D ResNet architectures was proposed. This study proved that the direct imagery-based deep learning model yielded competitive results, and both 2D and 3D ResNet models were capable of explaining nearly 90% of the variance observed in field-scale yield. Meng, S. Y. et al. pointed out that the acquisition of multitemporal satellite imagery was difficult and was even impossible in some specific regions during cloudy seasons. Therefore, they paid special attention to DL-based large-area crop mapping using one-shot hyperspectral spaceborne imagery. The hyperspectral imagery data employed in this study were collected by the OHS-2A satellite, which covers 150×150 km with a spatial resolution of 10 m and contains 256 bands with a spectral resolution of 2.5 nm, ranging from 400 to 1000 nm. The training and test samples, in areas of interest (AOI) 1, numbered 1635 and 1306 pixels, respectively. The AOI 2 training and test samples numbered 9889 and 10,269 pixels, respectively. Three types of CNN models, 1D CNN, 2D CNN, and 3D CNN were introduced to extract the spatial distribution features. As a result, the 3D CNN model achieved the best performance in terms of classification accuracy. Additionally, as compared with the mono-temporal or multitemporal MSI-based crop mapping, the classification accuracy when using hyperspectral satellite images was found to be higher than 94% [93].

It is worth mentioning that the quality of the images obtained from spectral cameras, including RGB, multispectral, and hyperspectral cameras, are sensitive to weather conditions. Synthetic aperture radar (SAR) imagery has become an essential complementary technique for agricultural remote sensing, given its unique characteristics, such as all-time and all-weather sensing [94,95]. However, the backscatter noise of the radar makes the interpretation of SAR imagery quite difficult. To deal with this problem, some DL-based frameworks have been proposed for object detection in SAR images [96,97]. For example, Zhang, T. W. et al. introduced a deep learning framework called Deep SAR-Net to determine the surfaces and objects with an average overall accuracy of 92.94 \pm 1.05 [97]. To further improve upon the representation ability of DL models, the concept of multi-sensor information fusion is introduced into deep learning. [98]. Specifically, the fusion of SAR

and optical imagery can contribute to better spatial and spectral information extraction [99]. For instance, Adrian, J. et al. investigated deep learning and machine learning-based crop type mapping by fusing multitemporal Sentinel-1 SAR and Sentinel-2 optical data [100]. Eight SAR ($136 \times 312 \times 10$) images and eight ($136 \times 312 \times 10$) optical images were used to generate eight SAR/optical fused images. To improve the training performance, each fused image was broken into 100 pixel patches ($40 \times 40 \times 10$). The mapping accuracies of 2D U-Net, 3D U-Net, SegNet, and Random Forest were evaluated by this dataset. The test results demonstrated that the fusion of multitemporal SAR and optical data yielded higher crop type mapping accuracy as compared with standalone multitemporal SAR or optical data, and 3D outperformed other methods as the corresponding accuracy reached 0.941. Due to the specific characteristics of optical and SAR sensors, it is challenging to model the relationship between the SAR and optical data. To solve this problem, Zhao, W. Z. et al. extended the conventional CNN-RNN to build a Multi CNN Sequence to Sequence (MCNN-Seq) model [101]. In this study, 31 Sentinel-1 SAR images and 73 Sentinel-2 optical images were collected from 5 January 2018 to 27 December 2018 and from 1 January 2018 to 27 December 2018, respectively. Trained by these data, the MCNN-Seq model could be used to predict optical time series using SAR data when the optical data was missing.

Given the coarse resolution and revisit period of satellite imagery, multitemporal, large-volume, and high-quality datasets are difficult to obtain, which impedes the application of deep learning in satellite-borne remote sensing. Consequently, DL-based satellite-borne remote sensing is mainly used for macroscopic observations at the field scale, such as land cover mapping, crop-background discrimination, and crop type classification [102,103]. This information is useful for large-scale yield prediction and nature disaster assessment; however, when it comes to small-scale detailed observations on crop growth, it is usually insufficient.

3. Transfer Learning (TL) in PA

The most significant limitation for the application of deep learning in precision agriculture is the lack of datasets [30]. Despite the fact that few public datasets consisting of specific plants have been created by agricultural communities, when the variability of environmental conditions and the biodiversity of agricultural ecosystems are taken into consideration, there are far from enough to relieve the data dependency of the supervised learning algorithms [25,37]. Transfer learning (TL) in the context of machine learning is a technique that aims to transfer the knowledge gained from source tasks, where large amounts of labeled datasets are available, to a new or similar target task in which training data are scarce [26]. TL can provide a better, cheaper, and faster solution to the issue of limited training data availability, especially, when the source and target domains have great similarities [22].

Based on our knowledge, we can broadly categorize TL into the following four strategies: (1) the pretrained architecture and parameters can be directly transferred into new tasks without additional retraining, (2) later layers of the neural network can be fine-tuned with an unseen dataset while freezing the other layers, (3) the entire model is fine-tuned with a new dataset when it is abundant enough, (4) the architecture and weights of pretrained models can be fine-tuned simultaneously. Two factors that specifically influence the transferability and suitability of the TL approach are the capacity of the dataset in the target domain and the similarity between the source and target tasks [21]. Generally, the more similar the new task is to the original task and the larger the new dataset, the higher the training speed and detection accuracy of transfer learning are. In PA, TL with pretrained backbones has been applied in plant identification, weed segmentation, disease detection, etc. In order to alleviate dependence on manpower, time, and expertise in deep neural network-based plant species identification, Kaya, A. et al. (2019) analyzed the effect of four different TL strategies (instance-transfer, feature-representation-transfer, parameter-transfer, and relational-knowledge-transfer) for plant species classification with four public datasets (Flavia, Swedish Leaf datasets, UCI Leaf dataset, and PlantVillage

dataset) [104]. This work proved that TL could provide important benefits for automated plant identification. In addition, TL was successfully utilized to locate plant centers with limited ground truth data by Cai, E. et al. (2020). They also concluded that TL was not effective if the distribution of the source domain was significantly divergent from the target domain [105]. Many well-known pretrained models, such as AlexNet, VGGNet, and GoogLeNet, can be incorporated into TL as backbones. Aimed at weed identification through transfer learning, Espejo-Garcia, B. et al. (2020) fine-tuned some state-of-the-art deep neural networks (Xception, Inception-Resnet, VGGNets, MobileNet, and DenseNet) and replaced the end fully connected network with other machine learning classifiers (logistic regression, support vector machines, and gradient boosting). A dataset that contained 504 RGB images of crops and weeds was established by taking images, weekly in May and June 2019 from 8:00 to 10:00, with a Nikon D700 camera at an approximate height of 1 m, at three different farms. Moreover, during all the image acquisition sessions, the illumination conditions were characterized as clear. The evaluation experiments revealed that the combination of fine-tuned DenseNet and SVM obtained a promising performance, with an F1 score of 99.29%, which proved their hypothesis that a specific combination of fine-tuning and classifier replacement may outperform fine-tuning only [28,106]. To determine if it was possible to transfer the knowledge of a network trained on a given crop to another crop type, and then to reduce the retraining time and efforts, Bosilj, P. et al. (2020) investigated the pixel-wise classification performance of an encoder-decoder CNN model that had been trained on three different datasets by means of TL. Three datasets, i.e., Sugar Beets 2016, Carrots 2017, and Onions 2017, were used. Specifically, Sugar Beet 2016 contained 283 images (1296×966) collected by the BoniRob agricultural robot with a JAI 130-GE camera in R, G, B, and NIR bands under controlled lighting conditions. Carrots 2017 covered 20 RGB and NIR images (2464 \times 2056) captured by two Teledyne DALSA Genie Nano cameras at a spatial resolution of 2.5 px/mm. Onions 2017 included 20 images with the same acquisition setup as Carrots 2017. This study found that transfer learning between crop types was not only helpful for the semantic segmentation of crops versus weeds, but also had the potential to reduce the training time up to 80% [107]. Concentrated on the semantic segmentation of oilseed rape in a field with high weed pressure, Abdalla, A. et al. (2019) evaluated three different TL methods using a VGG16-based encoder model and compared the differences in their performances with the deeper version of the VGG19 network. A tripod stabilized near-surface portable digital camera with focal length of 4 mm, ISO speed of 90, sensor size of 6.18×4.65 mm, aperture of 3.43 mm, and angle of view of 53.3°, was employed to capture images at a height of 1.2 m from the ground. A total of 400 RGB images (3888×5184) of oilseed rape crops along with their corresponding pixel-level annotations labeled by MATLAB 2018a Toolbox imageLabeler were used in this study, and random left/right reflection and random X/Y translation of -/+10 pixels were used as data augmentation techniques. The highest accuracy of 96% was achieved by the VGG16-based encoder, in which the fine-tuned model was only used for feature extraction and where the segmentation was performed using shallow machine learning classifiers [108]. In a similar work, Suh, H. K. et al. (2018) evaluated three different TL procedures for the classification of sugar beet and volunteer potato and assessed their performances amongst six CNN-based network architectures [109]. Apart from crop-weed classification, TL has also been used for crop disease detection. For instance, Chen, J. et al. (2020) proposed a novel convolutional neural network called INC-VGGN, which was developed by combining the VGGNet architecture with Inception modules and deployed the transfer learning strategy during the training process; 500 diseased rice and 466 diseased maize images captured under natural illumination intensities and clutter field background conditions were used in this study. All these images were adjusted to 224×224 pixels for training the deep leaning model. The results demonstrated that the proposed model was able to achieve an average accuracy of 92.00% for the classification of rice disease [29].

4. Few-Shot Learning (FSL) in PA

Although TL algorithms can reduce the dependence on datasets, large datasets with reliable annotations are still required. Instead of blindly expanding the capacity of a dataset, a novel deep learning method, namely one-shot or few-shot learning [27,110], has been proposed by mimicking human-level concept learning [111]. The most significant feature of FSL is "learning-to-learn" which means only a few labeled examples are needed to learn new concepts [112]. Over the past few years, many FSL-based models and algorithms have been extended to PA and have achieved great success.

Based on Siamese networks [113], Argüeso, D. et al. (2020) designed a few-shot learning (FSL) approach for plant disease classification in which three training architectures (a baseline fine-tunning Inception V3 network, a Siamese network with two subnets and contrastive loss, and a Siamese network with three subnets and triplet loss) with transfer learning were used [31]. One of the most significant results of this work was that when using the Siamese network with triplet loss and as few as 80 images per class, the accuracy was 90.0%, whereas when all 4421 training samples (six classes) from the target source were used by the baseline fine-tunning network, the accuracy was 94.0%. This means that FSL could reduce the training dataset size by 89.1% while only incurring a 4% loss in accuracy. This study preliminarily verified the feasibility of developing plant species and disease classification algorithms using very few labeled training data. In order to improve the deficiency in existing crop pest datasets, Li, Y. et al. (2020) introduced a few-shot learning method for the recognition of cotton pests [114]. The basic framework of the prototypical network was trained with triplet loss to extract embeddings. Two datasets were engaged in evaluating the effectiveness and feasibility of the few-shot model. The first dataset that contained 50 classes (10 samples for each class) of cotton pests was collected from the National Bureau of Agricultural Insect Resources (NBAIR). The other was a dataset with the natural scenes. The Euclidean distance was used in the following few-shot recognition workflow. In addition to accuracy performance, this work also concentrated on the flexibility and utility of the proposed approach, which were more important for the deployment of an embedded system. The testing accuracy exceeding 95% and a running speed that reached two frames per second on the embedded terminal made it more practical for use in a DL-based cotton pests recognition model in agricultural applications.

Another approach to improve the performance of FSL is the augmentation of the dataset by generating pseudo samples. In a study by Zhong, F. et al. (2020), a generative model using conditional adversarial autoencoders (CAAE) for the zero- and few-shot recognition of Citrus aurantium L. diseases was introduced. The authors built a dataset enveloping 1160 images of 9 types of Citrus aurantium L. diseases to evaluate the performance of proposed model. CAAE was used to enhance the dataset by generating synthetic samples. Extensive experiments indicated that this model performed well in recognizing Citrus aurantium L. diseases with only a few or even zero training samples [32]. Hu, G. et al. (2019) presented a low-shot learning method for the identification of tea leaf diseases. In this study, a conditional deep convolutional generative adversarial network (C-DCGAN) was employed to augment training samples, and the SVM classifier was used to segment disease spots from raw images. Thereafter, a deep learning network, VGG16, was trained by the augmented dataset. As a result, an average identification accuracy of 90% was achieved by combining traditional machine learning methods and deep learning methods [115].

As compared with RGB images, multispectral and hyperspectral images contain more subtle spectral-spatial agricultural information. However, the collection and annotation of abundant high-dimensional spectral images are still a bottleneck for supervised learning. To address this challenge, Liu, B., Gao, K., and their colleagues transferred deep few-shot learning (DFSL) technology into hyperspectral image classification where annotations have been scarce. For example, in 2019, they proposed a DFSL model to tackle the problem of small sample size [116]. In this work, a deep residual 3D CNN (D-Res-3-D CNN) was used to parameterize the metric space (Euclidean distance), and then to extract the spectral-spatial features. Later, in 2020, they added a relation network [117] and meta-learning [118]

in DFSL to improve the model's generalization ability and classification accuracy [119]. Although their study was not specific to the agricultural domain, it demonstrated the effectiveness of DFSL in hyperspectral image classification.

5. Discussion

Climate change, population growth, population aging, and the COVID-19 pandemic have posed significant threats to global agricultural development and food security [1,2,4]. Agricultural communities are committed to boosting food production and mitigating the risk of the global food crisis by introducing cutting-edge technologies, such as genomes, robots, and artificial intelligence into precision agriculture. Specifically, sensing technologies and autonomous systems are accelerating the digitalization of modern agriculture [5,6]. In this process, digital images with various modalities are becoming the most common type of data. Agricultural informatization can help farmers to observe farmland more comprehensively, and then carry out specific management. With the prevalence of satelliteborne remote sensing, aerial remote/proximal sensing, and field-robot-borne proximal sensing in agriculture, the volume of collected images has exploded. A critical issue is a method to quickly extract the most valuable details from these images. Traditionally, expert knowledge and profound experience are required to decode this information. However, it is time-consuming and labor-intensive and is usually prone to subjectivity. Over the past few years, we have witnessed significant advancements in deep learning, which has shown powerful capabilities of processing large amounts of images [17]. In precision agriculture, deep learning is playing an increasing role in addressing practical problems, such as crop identification, crop-weed classification, and stress monitoring [17,20,21,24]. In this section, we try to discuss the current situations, challenges, and suggestions regarding DL in PA from the following three aspects: datasets, deep learning algorithms and neural networks, and computational capacity.

5.1. Datasets

High-volume and high-quality agricultural datasets are the prerequisite for the deployment of DL in PA. Recently, Lu, Y. and Young, S. carried out a survey that discussed publicly available datasets for agricultural computer vision tasks [25]. This work covered a total of 34 datasets, in which 15 were for weed control, 10 were for fruit detection, and 9 were for miscellaneous applications. Each dataset was tailored for a specific task, and there were differences in the number of images, image types, image formats, and image sizes of different datasets. Some of these datasets were specially constructed for DL-based agricultural analytics. In our review, we briefly introduced the details of collection methods for datasets used in some representative studies. Among these datasets, some were collected under laboratory conditions [41,49,51,65], some were captured in the field [45,52,63,66,70,79], and others were established by combining these two methods [43,50,53]. Figure 2 shows some publicly available datasets for agriculture sensing at the leaf or canopy scale. CWFI, a benchmark dataset that consisted of 60 images (1296×966) of crops and weeds with annotations was proposed by Haug, S. and Ostemann, J. [120], These images were collected with an autonomous field robot in a carrot farmland. Although relatively small, it did stimulate subsequent research on crop/weed discrimination based on computer vision technology. PlantVillage, a popular dataset with 54,309 images (256×256) of healthy and diseased plants, was released in 2015 [41]. These images involved 14 crop species, 17 fungal diseases, 4 bacterial diseases, 2 mold diseases, 2 viral diseases, and 1 mite-induced disease. It is worth mentioning that an important limitation of this dataset was that the background was not an actual farmland scene but instead, was an ideal laboratory environment. Nevertheless, it has been widely used as a benchmark in DL-based disease diagnostics. Carrot Weed is a high-resolution (3264×2488) dataset for weed detection [121]. The dataset consists of 39 RGB images that were captured by cell phones from approximate heights of 1 m under natural variable light conditions. Pixel-level annotations for the crop, weed, and soil background are available. Wiesner-Hanks, T. et al. constructed an image dataset for

phenotyping Northern Leaf Blight (NLB) in maize fields [122]. It comprised 18,222 maize plant images with more than 100,000 NLB lesions. With the help of this dataset, the authors realized millimeter-level plant disease detection through deep learning [123]. DeepWeeds, a multiclass weed species image dataset for deep learning-based classification and agricultural robotic-based weed control, was proposed by Olsen et al. [68]. This publicly available dataset covers 17,509 in situ RGB images (256×256) of 8 weed species that are native to Australia. Recently, Liu et al. published the large-scale benchmark plant disease dataset 271 (PDD271) with 220,592 plant leaf images belonging to 271 plant disease categories for plant disease recognition [124]. The whole data construction process took about 2 years.



Figure 2. Some examples of public datasets for agriculture sensing at the leaf or canopy scale.
(a) CWFI (60 images) [120]. (b) PlantVillage (54,309 images) [39]. (c) Carrot-Weed (39 images) [121].
(d) MaizeDisease (18,222 images) [122]. (e) DeepWeeds (17,509 images) [68]. (f) PDD 217 (220,592 images) [124].

In addition to these close-range datasets, some large-area datasets of orthophotos have also appeared. For example, as shown in Figure 3, Chiu, M. T. et al. proposed Agriculture-Vision, a large aerial image database for agricultural pattern analysis [125]; 94,986 high-quality images with 4 channels (NIR, R, G, and B) captured from 3432 farmland samples were generated by cropping annotations with a window size of 512×512 pixels. This dataset was used for a pilot study of aerial agricultural semantic segmentation. Su, J. et al. proposed a multispectral (B, R, G, RedEdge, and NIR) and RGB dataset (1600) by slicing image tiles [256,256] from UAV-based orthomosaic imagery [126].





Based on the above analysis, we found some issues to be aware of:

- (i). Most datasets are built for specific tasks and are not general.
- (ii). The volume of the dataset is generally small, and only a small part can reach the volume of tens of thousands of images.
- (iii). The background of the images is usually simple, especially for datasets collected in a laboratory setting. This leads to insufficient generalization ability of deep learning in unstructured and complex farmland environments.
- (iv). Each dataset differs from the others in terms of camera configurations, employed platforms, collection periods, ground sampling resolutions, imagery types, and so on. This results in insufficient adaptability of the dataset across different tasks.
- (v). In some datasets, images from different classes are imbalanced, which can lead to biased accuracy.
- (vi). Most datasets have varying degrees of temporal and spatial limitations, which means that images in datasets only contain information about short-term local regions.
- (vii). Vertically downward is the most common method, which lacks the representation of multiple angles of the same target.
- (viii).Handheld cameras, terrestrial agricultural robots, and UAVs are the most common platforms for collecting farm imagery, however, their combination is rarely used.

Generally, the larger the volume of the dataset, the higher the quality, the less likely it is for overfitting to occur during deep learning network training, and the higher the accuracy. However, considering the complexity of farmland environments, the heterogeneity of crops, diseases, pests, and weeds, as well as the diversity of information collection equipment, these datasets are far from what is actually needed, both in quantity and quality. Nowadays, the collection, preprocessing, and annotation of a high-quality dataset requires extensive labor, as well as financial and time inputs which makes it difficult or even impossible to construct a general dataset that is robust for multiple similar agricultural tasks rather than a single one. To deal with this issue, we provide some suggestions for agricultural dataset acquisition:

- (i). Collection methods. In order to ensure the generalization ability and robustness of deep learning algorithms in unstructured and complex farmland environments, if possible, images should be collected in real farmlands, rather than in controllable laboratory environments. Before image acquisition, there should be detailed plans for the observation object, acquisition location, acquisition period, acquisition time, acquisition equipment, camera parameters, etc., and efforts should be made to keep these consistent throughout the entire dataset construction process. Complex backgrounds, natural lighting, observation angles, perspective conditions, etc. should also be taken into consideration, and these should be as close to the real application scenario as possible. Combining high-performance cameras with mobile platforms such as handheld devices, field robots, and UAVs to collect high-quality images at multiple spatial and time scales is necessary for the construction of future agricultural datasets. For example, multiple acquisition platforms can be used to collect information on the same observation object at multiple spatial scales and perspectives at the same time. In addition, soil, climate, and environmental information should also be accurately recorded during image acquisition, as this information has the potential to further improve model performance.
- (ii). Raw data processing. The selection, classification, preprocessing, and labeling of raw images should be done under the guidance of experienced agricultural scientists or farmers. It is worth noting that the data balance between different categories should be given special attention when constructing a dataset. Some data augmentation methods can be used to further increase the size of a dataset. For example, a dataset can be augmented by filtering images from existing public datasets that are similar to the target task. Some conventional data augmentation methods such as rotation, mirroring, scaling, and adding noise, are also effective for augmenting datasets. In addition, pseudo samples generated by artificial intelligence methods are also effective

in situations where it is difficult to construct large-scale datasets. Usually, limited by the neural network model and computing resources, the original image needs to be cropped or scaled to a specific size before it can be used as the input of the network. During this process, the integrity of the image details should be maintained as much as possible.

- (iii). Temporal resolution. Throughout an entire crop growth cycle, the farmland environment changes significantly, and the information of the same growth stage in different planting seasons is not consistent. Therefore, the temporal resolution of the dataset also needs attention. Some datasets have taken the changes of physical morphology and physiological properties of crops and weeds during various growth stages into account. Meanwhile, the acquisition period of some datasets spanned a few consecutive seasons. Further refining the temporal resolution of the dataset acquisition period is very important to improve the robustness of the deep learning model algorithm.
- (iv). Spatial and spectral resolution. Image diversity plays a critical role in increasing spatial and spectral resolutions of the dataset. RGB cameras, the most common sensors, can yield high spatial resolution images with visible spectrum bands. RGB images contain rich features of color, texture, and profile that are essential for DL-based classification, detection, segmentation, and tracking. However, the quality of RGB images is very sensitive to light and can only present information in the visible spectrum. Recently, multispectral and thermal imagery has been playing an increasing role in agricultural sensing, since the information underlying in invisible spectrum significantly helps with the detection of early crop deficits. Satellite-borne SAR and optical images have also been broadly used for large-scale farmland mapping. RGB, spectral, thermal, and SAR cameras provide rich images at various spatial and spectral resolutions. We suggest that future datasets should be constructed with a combination of multiple types of cameras to compensate for the resolution limitations of a single type of sensor.
- (v). Sharing and maintenance. We found that datasets built by different teams for the same target task vary widely. In order to improve the generality and utilization efficiency of datasets, cross-regional multilateral cooperation among different research teams is needed. For example, the species of rice, weeds, and diseases vary in different countries and regions. If the collection and sharing methods of datasets can achieve some kind of consensus and cooperation among multiple research teams, the volume and quality of the corresponding datasets will be greatly improved. In addition, the maintenance and supplementation of datasets is also related to the long-term sustainability of scientific research. Although some researchers publicly shared their datasets at the time of paper publication, parts of the datasets had either been lost over time or become obsolete due to poor maintenance, which hindered further research.

5.2. Deep Learning Algorithms and Neural Networks

Another factor that limits the accuracy and speed of agricultural information interpretation is the DL algorithm and network architecture. Following the timeline, Table 1 lists some representative studies that are related to DL applications throughout the whole crop growth cycle, such as land mapping, biotic/abiotic stress identification, classification and quantification, and yield prediction. The applications of DL in PA have developed rapidly since 2016. Among the many published scientific papers, a variety of state-of-the-art DL frameworks and algorithms, such as AlexNet, VGG, SegNet, ResNet, Inception, U-Net, and Mask RCNN, have been proposed for agricultural image classification, detection, and segmentation.

Categories	Scale	Application	Dataset	Network Architecture	Performance	Year	Refs
CNN-SL	Leaf	Disease detection	PlantVillage	AlexNet, GoogLeNet	CA, 99.35%	2016	[42]
CNN-SL	Canopy	Plant species classification	22 Weed and crop species	Customized	CA, 33–98%	2016	[61]
CNN-SL	Canopy	Weed detection	Grass-Broadleaf	AlexNet	CA, 98%	2017	[70]
CNN-SL	Leaf	Disease detection	PlantVillage	AlexNet, AlexNetOWTBn, GoogLeNet, Overfeat, VGG	SR, 99.53%	2018	[43]
CNN-SL	Canopy	Joint stem detection and crop/weed classification	Sugar beet and dicot weeds	FCN architecture Stem-seg-S	mAP, 85.4% (stem detection) 69.7% (segmentation)	2018 2020	[63,64]
CNN-SL	Canopy	Crop/weed classification	Plant seedlings dataset (3 crops and 9 weeds)	AgroAVNET	CA, 98.23%	2018	[65]
TL	Canopy	Crop/weed classification	Sugar beet and volunteer potato	AlexNet, VGG-19, GoogLeNet, ResNet-50, ResNet-101, Inception-v3.	Highest CA, 96% (VGG-19)	2018	[109]
CNN-SL	Leaf	Disease classification	17 Diseases on 5 crops	ResNet50 variants	BA, 98%	2019	[45]
CNN-SL	Leaf	Disease identification	Charcoal rot disease	3D DCNN	CA, 95.73%	2019	[49]
CNN-SL	Leaf	Disease severity detection	Sugar beet leaf images	Updated Faster R-CNN	CA, 95.48%	2019	[50]
CNN-SL	Canopy	Weed specie classification	DeepWeeds	Inception-v3 and ResNet-50	Highest CA, 95.7%	2019	[68]
CNN-SL	Field	Wheat and barley yield prediction	RGB and NDVI images	Customized	Lowest MAPE, 8.8	2019	[86]
TL	Leaf	Plant classification	Flavia, Swedish Leaf, UCI Leaf, PlantVillage	AlexNet/VGG-16 with LSTM	Highest CA, 99.11% (Swedish Leaf)	2019	[104]
TL	Canopy	Crop/weed semantic segmentation	Oilseed rape images	VGG-based encoder	Highest CA, 96%	2019	[108]
FSL	Leaf	Disease identification	Tea leaf's disease images	C-DCGAN-VGG16	Average identification accuracy, 90%	2019	[116]
CNN-SL	Leaf	Pest recognition	10 Common species of crop pest	VGG-16, VGG-19, ResNet50, ResNet152, GoogLeNet	CA, 96.67%	2020	[53]
CNN-SL	Leaf	Pest detection and recognition	In-field Pest in Food Crop (IPFC)	DeepPest	mAP, 74.3%	2020	[55]

Table 1. Examples of DL applications in PA following the timeline.

|--|

Categories	Scale	Application	Dataset	Network Architecture	Performance	Year	Refs
CNN-SL	Canopy	Weed and crop recognition	Four different weed datasets	GCN-ResNet-101	Highest CA, 99.37%	2020	[66]
CNN-SL	Canopy	Weed recognition	DeepWeeds	Graph Weeds Net (GWN)	Top-1 accuracy: 98.1%	2020	[69]
CNN-SL	Field	Water stress identification	Maize, okra, and soybean (1200)	AlexNet, GoogLeNet, Inception V3	CA (GoogLeNet), 98.30, 97.50, 94.16%	2020	[74]
CNN-SL	Field	Salt-stress phenotyping	Okra hyperspectral images	Encoder-Decoder	IoU, 0.94 (plant segmentation) BD, 85.4 (leaf segmentation)	2020	[75]
CNN-SL	Field	Growth stage estimation	Wheat and barley images	5-layer ConvNet, VGG19	Accuracy, 91.1–94.2% (ConvNet), 99.7–100% (VGG19)	2020	[79]
CNN-SL	Field	Nutrient Status Diagnosis	Oilseed rape RGB images	CNN-LSTM	Highest CA, 95% (Inceptionv3-LSTM)	2020	[80]
CNN-SL	Field	Wheat lodging extraction	RGB, multispectral images	TL and Deeplab v3+	Highest DC, 0.90 (early maturity with RGB) and 0.93 (early maturity with multispectral)	2020	[81]
CNN-SL	Field	Cotton yield predication	RGB images	Faster R-CNN, SSD, SSDLITE	Mean errors of 17.86%	2020	[85]
CNN-SL	Land	Crop monitoring	Satellite SAR and optical images	MCNN-Seq	Highest R ² , 0.9824	2020	[101]
TL	Canopy	Plant Center Localization	Maize orthomosaic images	ResNet encoder modified U-Net	Precision, 84.3%; Recall, 98.8%; F1 Score, 91.0%	2020	[105]
TL	Canopy	Crop/weed identification	Two crops and two weeds	Xception, Inception-Resnet, VGNets, MobileNet, DenseNet	F1 Score, 99.29%	2020	[106]
TL	Canopy	Crop/weed semantic segmentation	Sugar Beets 2016, Carrots 2017, Onions 2017	SegNet	Training times reduced 80%	2020	[107]
TL	Leaf	Plant disease identification	Rice/maize plant images	INC-VGGN	Average accuracy, 92%	2020	[29]
FSL	Leaf	Plant disease classification	PlantVillage	Inception V3-Siamese network	Accuracy above 90% and 89.1% reduction in training data	2020	[31]

Categories	Scale	Application	Dataset	Network Architecture	Performance	Year	Refs
FSL	Leaf	Pest recognition	Two Cotton dataset	CNN- prototypical network	Testing accuracy, 95.4% and 96.2%	2020	[114]
FSL	Leaf	Disease recognition	Citrus aurantium L. diseases images	CAAE	Harmonic mean accuracy of 53.4%	2020	[32]
FSL	Leaf	Pest classification	PlantVillage	Shallow CNN	Varies under different conditions	2021	[33]
CNN-SL	Leaf	Disease severity quantification	NLB-infected maize images	Cascaded MRCNN	Correlation 73%, efficiency of 5 fps	2021	[52]
CNN-SL	Land	Yield predication	Satellite multispectral images	2D and 3D ResNet	Nearly 90%	2021	[92]
CNN-SL	Land	Crop mapping	Satellite hyperspectral images	1D/2D/3D CNN	More than 94% (3D CNN)	2021	[93]
CNN-SL	Land	Land mapping	SAR images	Deep SAR-Net	Accuracy, 92.94 \pm 1.05	2021	[98]
CNN-SL	Land	Crop mapping	SAR and optical images	3D U-Net	Overall accuracy, 0.941	2021	[100]

Abbreviations: CNN-SL, convolutional neural network-based supervised learning; TL, transfer learning; FSL, few-shot learning; CA, classification accuracy; SR, success rate; BA, balanced accuracy; MSE, mean square error; mAP, mean average precision; IoU, intersection over union; BD, best dice; DC, dice coefficients; MAPE, mean absolute percentage error.

However, some issues and limitations need to be pointed out:

- (i). We found that most of the DL-based methods applied in PA only used simple algorithms and network structures. The main reason is that the combination of deep learning and precision agriculture is still in its infancy. Another reason for this phenomenon is the isolation of interdisciplinary knowledge between agricultural scientists and computer scientists.
- (ii). Although many of the deep learning algorithms listed in Table 1 achieved an accuracy of more than 90%, it should be made clear that these results are limited to specific datasets. When the trained models are deployed to other datasets or a real farmland environment, the accuracy and speed of these networks are usually lower than the benchmarks. The main reason for this is that the volume, quality, and complexity of agricultural datasets are still quite different from actual farmland environments.
- (iii). Some novel strategies, such as transfer learning [104–109], few-shot learning [31,32,114,115], graph convolutional networks (GCN) [66], generative adversarial networks (GAN) [127], and semi-supervised learning [128] have been proposed to reduce the dependency of DL models on the agricultural datasets. However, their performance has not been fully released.

Recently, in a small number of studies, deep learning algorithms and neural network architectures have been improved and optimized for the characteristics of agricultural applications. For example, some research has focused on the optimization of inputs of DL models. Su, J. et al. formulated the monitoring and quantifying yellow rust as a supervised pixelwise multiclass classification problem and introduced U-Net to classify the ruse, healthy, and backgrounds in wheat fields [126]. Data input optimization was the highlight of this work. Benefiting from the rich spectral information collected by multispectral and RGB cameras, they used spectral-spatial information from various spectral bands and generated vegetation index as the U-Net input instead of the commonly used RGB images. They claimed that the introduction of the RedEdge and NIR bands could improve segmentation performance over visible RGB images. A similar work was conducted by Maimaitijiang, M. et al. in 2020 [129]. In this study, UAV-based multimodal data fusion using RGB, multispectral, and thermal sensors to predict soybean yield were comprehensively investigated. According to different information fusion strategies, they proposed two deep neural network structures, called DNN-F1 and DNN-F2, respectively. DNN-F1 refers to input-level feature fusion, and DNN-F2 refers to intermediated-level feature fusion. The results also indicated that DNN-based multimodal data fusion improved model accuracy. Other studies have focused on improving DL algorithms and frameworks. Sa, I. et al. presented WeedNet and WeedMap, the modified versions of the SegNet architecture, for the large-scale dense semantic segmentation of weeds with aerial images [130,131]. Specifically, they replaced the origin encoder with a modified VGG16 architecture and amended the decoder accordingly. In order to balance the accuracy and speed of deep learning algorithms, Jiao, L. et al. designed an anchor-free convolutional neural network (AF-RCNN) for multiclass agricultural pest detection [132].

To advance the application of deep learning in precision agriculture, we recommend:

- (i). More extensive optimization of deep learning algorithms and neural networks based on the agricultural characteristics that are needed. Research has demonstrated the effectiveness of some biomimetic mechanisms, such as visual attention mechanism and long short-term memory mechanism, in image processing. Therefore, they should also be introduced into model optimization for agricultural applications.
- (ii). As the spatial, temporal, and spectral resolution of the dataset increases, there is a strong need for neural networks that can take data cubes as input.
- (iii). Efforts in this area should not be limited to only data-driven deep learning. Metalearning, which means learning-to-learn, is also a research direction. It has been proven that transfer learning, few-shot learning, and semi-supervised learning have great potential to relieve dependency on datasets, especially in situations where datasets are difficult to obtain. However, they inevitably cause a decrease in accuracy.

22 of 27

In some cases, this is acceptable and in others it is not. Therefore, more efforts should be made to address this contradiction or find a trade-off between dataset and performance.

- (iv). More attention should be given to semi-supervised learning. For example, the recently emerged GCN is a successful semi-supervised learning approach in which a large amount of unlabeled data can be leveraged with typically a small amount of labeled data.
- (v). Interdisciplinarity is a necessary way to break down the intellectual segregation between agricultural scientists and computer scientists. The close communication and cooperation between the two parties is of great significance for deepening the application of deep learning in precision agriculture.

5.3. Computational Capacity

Another bottleneck that limits the deployment of DL in PA is computational capacity. With an increase in the number of datasets and the deepening of deep learning neural networks, model training requires more and more computing resources. The performance improvement of the central processing unit (CPU) and the graphics processing unit (GPU) is a very important accelerator for the widespread popularity of deep learning. Additionally, Cloud computing services, such as the Google Cloud Platform, which are provided by commercial companies, have accelerated the development of DL. However, due to strict computation capacity requirements, the current DL in PA is usually offline, which means that the collection and analysis of farmland images is asynchronous. This does not comply with the principle of PA, which aims to carry out detection and treatment synchronously. For example, UAVs have been widely used for agricultural information collection and agrichemical spraying. Additionally, UAV-based spraying relies on a prescription generated by data analysis of farmland information. Limited by the computing power of an airborne processor, the acquisition of farmland information based on deep learning cannot be performed in the air in real time. Consequently, it is difficult for present agricultural UAVs to carry out precision treatment by running DL-based agri-analytics onboard. Fortunately, this issue has attracted some attention from companies. For instance, some embedded systems, such as Jetson TX2 and AGX Xavier, have been developed specifically for autonomous robots by NVIDIA (https://www.nvidia.cn/autonomous-machines/embedded-systems/ accessed on 10 December 2021). More high-performance, low-power, and lightweight processors may be available in the future.

6. Conclusions

Precision agriculture aims to increase crop yield while reducing or at least maintaining resource inputs by introducing cutting-edge technology. Deep learning has great potential to advance the evolution of precision agriculture. This work paid special attention to the application of deep learning in precision agriculture. Specifically, our main focus was on three kinds of deep learning methods, namely CNN-SL, TL, and FSL. From a spatial perspective, we defined four scales, the leaf scale, canopy scale, field scale, and land scale. Some studies that have discussed their application in multiscale agricultural sensing were reviewed. Challenges and opportunities for DL in PA were discussed. We hope that this work will attract the attention of agricultural communities and that it will promote more relevant research involving DL in PA.

Author Contributions: D.W. and X.W. designed the framework of this review; D.W., W.C., F.Z. and Z.L. participated in the survey and analysis of relevant papers; D.W. and F.Z. wrote the first draft of the manuscript; Z.L., S.X. and W.C. revised and edited the final version of the manuscript; D.W. and S.X. are responsible for funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (32001424), the Shenzhen Science and Technology Program (JCYJ20210324102401005), and the China Postdoctoral Science Foundation (2020M682989).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- 1. World Bank. World Development Report 2008: Agriculture for Development; The World Bank: Washington, DC, USA, 2007. [CrossRef]
- 2. Food and Agriculture Organization of the United Nations. *Agriculture and Climate Change: Challenges and Opportunities at the Global and Local Level: Collaboration on Climate-Smart Agriculture;* FAO: Rome, Italy, 2019.
- Fedoroff, N.V.; Battisti, D.S.; Beachy, R.N.; Cooper, P.J.M.; Fischhoff, D.A.; Hodges, C.N.; Knauf, V.C.; Lobell, D.; Mazur, B.J.; Molden, D.; et al. Radically rethinking agriculture for the 21st century. *Science* 2010, 327, 833–834. [CrossRef]
- 4. Laborde, D.; Martin, W.; Swinnen, J.; Vos, R. COVID-19 risks to global food security. Science 2020, 369, 500–502. [CrossRef]
- 5. King, A. Technology: The future of agriculture. *Nature* **2017**, 544, S21–S23. [CrossRef]
- 6. Asseng, S.; Asche, F. Future farms without farmers. Sci. Robot. 2019, 4, eaaw1875. [CrossRef]
- 7. Gebbers, R.; Adamchuk, V.I. Precision agriculture and food security. *Science* 2010, 327, 828–831. [CrossRef]
- Penuelas, J.; Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* 1998, 3, 151–156. [CrossRef]
- 9. Burke, M.; Driscoll, A.; Lobell, D.B.; Ermon, S. Using satellite imagery to understand and promote sustainable development. *Science* **2021**, *371*, eabe8628. [CrossRef]
- 10. Maes, W.H.; Steppe, K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* **2019**, *24*, 152–164. [CrossRef]
- 11. Vougioukas, S.G. Agricultural robotics. Annu. Rev. Control Robot. Auton. Syst. 2019, 2, 365–392. [CrossRef]
- 12. Singh, A.; Ganapathysubramanian, B.; Singh, A.K.; Sarkar, S. Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* **2016**, *21*, 110–124. [CrossRef]
- Ma, C.; Zhang, H.H.; Wang, X.F. Machine learning for big data analytics in plants. *Trends Plant Sci.* 2014, 19, 798–808. [CrossRef] [PubMed]
- Li, D.L.; Li, C.; Yao, Y.; Li, M.; Liu, L. Modern imaging techniques in plant nutrition analysis: A review. Comput. Electron. Agric. 2020, 174, 105459. [CrossRef]
- 15. Patrício, D.I.; Rieder, R. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Comput. Electron. Agric.* **2018**, *153*, 69–81. [CrossRef]
- Xu, S.; Liu, J.; Yang, C.; Wu, X.; Xu, T. A learning-based stable servo control strategy using broad learning system applied for microrobotic control. *IEEE Trans. Cybern.* 2021, 1–11. [CrossRef]
- 17. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature. Nature 2015, 521, 436–444. [CrossRef]
- 18. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. Comput. Electron. Agric. 2018, 147, 70–90. [CrossRef]
- 19. Ma, L.; Liu, Y.; Zhang, X.L.; Ye, Y.X.; Yin, G.F.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm.* **2019**, *152*, 166–177. [CrossRef]
- Zhong, L.H.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* 2019, 221, 430–443. [CrossRef]
- 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]
- 22. Singh, A.; Jones, S.; Ganapathysubramanian, B.; Sarkar, S.; Mueller, D.; Sandhu, K.; Nagasubramanian, K. Challenges and opportunities in machine-augmented plant stress phenotyping. *Trends Plant Sci.* **2021**, *26*, 53–69. [CrossRef]
- 23. Wang, D.; Li, W.; Liu, X.; Li, N.; Zhang, C. UAV environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution. *Comput. Electron. Agric.* **2020**, *175*, 105523. [CrossRef]
- Hasan, A.S.M.M.; Sohel, F.; Diepeveen, D.; Laga, H.; Jones, M.G. A survey of deep learning techniques for weed detection from images. *Comput. Electron. Agric.* 2021, 184, 106067. [CrossRef]
- 25. Lu, Y.Z.; Young, S. A survey of public datasets for computer vision tasks in precision agriculture. *Comput. Electron. Agric.* 2020, 178, 105760. [CrossRef]
- 26. Pan, S.J.; Yang, Q. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 2009, 22, 1345–1359. [CrossRef]
- 27. Snell, J.; Swersky, K.; Zemel, R.S. Prototypical networks for few-shot learning. arXiv 2017, arXiv:1703.05175.
- Espejo-Garcia, B.; Mylonas, N.; Athanasakos, L.; Fountas, S.; Vasilakoglou, I. Towards weeds identification assistance through transfer learning. *Comput. Electron. Agric.* 2020, 171, 105306. [CrossRef]

- Chen, J.; Chen, J.X.; Zhang, D.F.; Sun, Y.D.; Nanehkaran, Y.A. Using deep transfer learning for image-based plant disease identification. *Comput. Electron. Agric.* 2020, 173, 105393. [CrossRef]
- 30. Barbedo, J.G.A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* 2018, 153, 46–53. [CrossRef]
- 31. Argüeso, D.; Picon, A.; Irusta, U.; Medela, A.; San-Emeterio, M.G.; Bereciartua, A.; Alvarez-Gila, A. Few-shot learning approach for plant disease classification using images taken in the field. *Comput. Electron. Agric.* **2020**, *175*, 105542. [CrossRef]
- 32. Zhong, F.M.; Chen, Z.K.; Zhang, Y.C.; Xia, F. Zero-and few-shot learning for diseases recognition of *Citrus aurantium* L. using conditional adversarial autoencoders. *Comput. Electron. Agric.* **2020**, *179*, 105828. [CrossRef]
- 33. Li, Y.; Yang, J. Meta-learning baselines and database for few-shot classification in agriculture. *Comput. Electron. Agric.* 2021, 182, 106055. [CrossRef]
- 34. Lee, S.H.; Chan, C.S.; Mayo, S.J.; Remagnino, P. How deep learning extracts and learns leaf features for plant classification. *Pattern Recognit.* **2017**, *71*, 1–13. [CrossRef]
- 35. Grinblat, G.L.; Uzal, L.C.; Larese, M.G.; Granitto, P.M. Deep learning for plant identification using vein morphological patterns. *Comput. Electron. Agric.* 2016, 127, 418–424. [CrossRef]
- 36. Barbedo, J.G. Factors influencing the use of deep learning for plant disease recognition. Biosyst. Eng. 2018, 172, 84–91. [CrossRef]
- Strange, R.N.; Scott, P.R. Plant disease: A threat to global food security. Annu. Rev. Phytopathol. 2015, 43, 83–116. [CrossRef] [PubMed]
- Noon, S.K.; Amjad, M.; Qureshi, M.A.; Mannan, A. Use of deep learning techniques for identification of plant leaf stresses: A review. Sustain. Comput. Inform. Syst. 2020, 28, 100443. [CrossRef]
- Zeng, W.H.; Li, M. Crop leaf disease recognition based on Self-Attention convolutional neural network. *Comput. Electron. Agric.* 2020, 172, 105341. [CrossRef]
- 40. Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* 2019, 161, 272–279. [CrossRef]
- Hughes, D.P.; Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv 2015, arXiv:1511.08060.
- 42. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 2016, 7, 1419. [CrossRef]
- 43. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, 145, 311–318. [CrossRef]
- 44. Lee, S.H.; Goëau, H.; Bonnet, P.; Joly, A. New perspectives on plant disease characterization based on deep learning. *Comput. Electron. Agric.* **2020**, 170, 105220. [CrossRef]
- Picon, A.; Seitz, M.; Alvarez-Gila, A.; Mohnke, P.; Ortiz-Barredo, A.; Echazarra, J. Crop conditional convolutional neural networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. *Comput. Electron. Agric.* 2019, 167, 105093. [CrossRef]
- 46. Picon, A.; Alvarez-Gila, A.; Seitz, M.; Ortiz-Barredo, A.; Echazarra, J.; Johannes, A. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Comput. Electron. Agric.* **2019**, *161*, 280–290. [CrossRef]
- 47. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [CrossRef]
- 48. Li, Y.; Zhang, H.K.; Shen, Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* **2017**, *9*, 67. [CrossRef]
- 49. Nagasubramanian, K.; Jones, S.; Singh, A.K.; Sarkar, S.; Singh, A.; Ganapathysubramanian, B. Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods* **2019**, *15*, 98. [CrossRef]
- 50. Ozguven, M.M.; Adem, K. Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Phys. A Stat. Mech. Its Appl.* **2019**, *535*, 122537. [CrossRef]
- Lin, K.; Gong, L.; Huang, Y.X.; Liu, C.L.; Pan, J.H. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. *Front. Plant Sci.* 2019, 10, 155. [CrossRef]
- 52. Garg, K.; Bhugra, S.; Lall, B. Automatic quantification of plant disease from field image data using deep learning. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 1965–1972.
- 53. Li, Y.F.; Wang, H.X.; Dang, L.M.; Sadeghi-Niaraki, A.; Moon, H. Crop pest recognition in natural scenes using convolutional neural networks. *Comput. Electron. Agric.* 2020, 169, 105174. [CrossRef]
- 54. Cheng, X.; Zhang, Y.H.; Chen, Y.Q.; Wu, Y.Z.; Yue, Y. Pest identification via deep residual learning in complex background. *Comput. Electron. Agric.* **2017**, *141*, 351–356. [CrossRef]
- 55. Wang, F.Y.; Wang, R.J.; Xie, C.J.; Yang, P.; Liu, L. Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition. *Comput. Electron. Agric.* **2020**, *169*, 105222. [CrossRef]
- 56. Wu, X.L.; Aravecchia, S.; Lottes, P.; Stachniss, C.; Pradalier, C. Robotic weed control using automated weed and crop classification. *J. Field Robot.* **2020**, *37*, 322–340. [CrossRef]
- 57. Gai, J.Y.; Tang, L.; Steward, B.L. Automated crop plant detection based on the fusion of color and depth images for robotic weed control. *J. Field Robot.* **2020**, *37*, 35–52. [CrossRef]

- 58. Dyrmann, M.; Christiansen, P.; Midtiby, H.S. Estimation of plant species by classifying plants and leaves in combination. *J. Field Robot.* **2018**, *35*, 202–212. [CrossRef]
- 59. Slaughter, D.C.; Giles, D.K.; Downey, D. Autonomous robotic weed control systems: A review. *Comput. Electron. Agric.* 2008, 61, 63–78. [CrossRef]
- 60. Knoll, F.J.; Czymmek, V.; Poczihoski, S.; Holtorf, T.; Hussmann, S. Improving efficiency of organic farming by using a deep learning classification approach. *Comput. Electron. Agric.* **2018**, 153, 347–356. [CrossRef]
- 61. Dyrmann, M.; Karstoft, H.; Midtiby, H.S. Plant species classification using deep convolutional neural network. *Biosyst. Eng.* **2016**, 151, 72–80. [CrossRef]
- 62. Lottes, P.; Hörferlin, M.; Sander, S.; Stachniss, C. Effective vision-based classification for separating sugar beets and weeds for precision farming. *J. Field Robot.* 2017, 34, 1160–1178. [CrossRef]
- 63. Lottes, P.; Behley, J.; Chebrolu, N.; Milioto, A.; Stachniss, C. Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *J. Field Robot.* **2020**, *37*, 20–34. [CrossRef]
- 64. Lottes, P.; Behley, J.; Milioto, A.; Stachniss, C. Fully convolutional networks with sequential information for robust crop and weed detection in precision farming. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2870–2877. [CrossRef]
- 65. Chavan, T.R.; Nandedkar, A.V. AgroAVNET for crops and weeds classification: A step forward in automatic farming. *Comput. Electron. Agric.* **2018**, 154, 361–372. [CrossRef]
- 66. Jiang, H.H.; Zhang, C.Y.; Qiao, Y.L.; Zhang, Z.; Zhang, W.J.; Song, C.Q. CNN feature based graph convolutional network for weed and crop recognition in smart farming. *Comput. Electron. Agric.* **2020**, *174*, 105450. [CrossRef]
- 67. Wang, A.; 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]
- Olsen, A.; Konovalov, D.A.; Philippa, B.; Ridd, P.; Wood, J.C.; Johns, J.; Banks, W.; Girgenti, B.; Kenny, O.; Whinney, J.; et al. DeepWeeds: A multiclass weed species image dataset for deep learning. *Sci. Rep.* 2019, *9*, 2058. [CrossRef]
- 69. Hu, K.; Coleman, G.; Zeng, S.; Wang, Z.; Walsh, M. Graph weeds net: A graph-based deep learning method for weed recognition. *Comput. Electron. Agric.* 2020, 174, 105520. [CrossRef]
- 70. dos Santos Ferreira, A.; Freitas, D.M.; da Silva, G.G.; Pistori, H.; Folhes, M.T. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* 2017, 143, 314–324. [CrossRef]
- dos Santos Ferreira, A.; Freitas, D.M.; da Silva, G.G.; Pistori, H.; Folhes, M.T. Unsupervised deep learning and semi-automatic data labeling in weed discrimination. *Comput. Electron. Agric.* 2019, 165, 104963. [CrossRef]
- 72. Mittler, R. Abiotic stress, the field environment and stress combination. *Trends Plant Sci.* 2009, 11, 15–19. [CrossRef]
- 73. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.C.; Jha, S.K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.* 2020, *21*, 1121–1155. [CrossRef]
- Chandel, N.S.; Chakraborty, S.K.; Rajwade, Y.A.; Dubey, K.; Tiwari, M.K.; Jat, D. Identifying crop water stress using deep learning models. *Neural Comput. Appl.* 2021, 33, 5353–5367. [CrossRef]
- 75. Feng, X.; Zhan, Y.; Wang, Q.; Yang, X.; Yu, C.; Wang, H.; Tang, Z.; Jiang, D.; Peng, C.; He, Y. Hyperspectral imaging combined with machine learning as a tool to obtain high-throughput plant salt-stress phenotyping. *Plant J.* 2020, 101, 1448–1461. [CrossRef] [PubMed]
- 76. Lee, K.J.; Lee, B.W. Estimation of rice growth and nitrogen nutrition status using color digital camera image analysis. *Eur. J. Agron.* **2013**, *48*, 57–65. [CrossRef]
- 77. Velumani, K.; Madec, S.; de Solan, B.; Lopez-Lozano, R.; Gillet, J.; Labrosse, J.; Jezequel, S.; Comar, A.; Baret, F. An automatic method based on daily in situ images and deep learning to date wheat heading stage. *Field Crops Res.* 2020, 252, 107793. [CrossRef]
- Barbedo, J.G.A. Detection of nutrition deficiencies in plants using proximal images and machine learning: A review. *Comput. Electron. Agric.* 2019, 162, 482–492. [CrossRef]
- Rasti, S.; Bleakley, C.J.; Silvestre, G.C.; Holden, N.M.; Langton, D.; O'Hare, G.M. Crop growth stage estimation prior to canopy closure using deep learning algorithms. *Neural Comput. Appl.* 2021, 33, 1733–1743. [CrossRef]
- Abdalla, A.; Cen, H.Y.; Wan, L.; Mehmood, K.; He, Y. Nutrient status diagnosis of infield oilseed rape via deep learning-enabled dynamic model. *IEEE Trans. Ind. Inform.* 2020, 17, 4379–4389. [CrossRef]
- Zhang, D.Y.; Ding, Y.; Chen, P.F.; Zhang, X.Q.; Pan, Z.G.; Liang, D. Automatic extraction of wheat lodging area based on transfer learning method and deeplabv3+ network. *Comput. Electron. Agric.* 2020, 179, 105845. [CrossRef]
- Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* 2018, 151, 61–69. [CrossRef]
- Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* 2020, 177, 105709. [CrossRef]
- 84. Barbosa, A.; Trevisan, R.; Hovakimyan, N.; Martin, N.F. Modeling yield response to crop management using convolutional neural networks. *Comput. Electron. Agric.* 2020, 170, 105197. [CrossRef]
- 85. Tedesco-Oliveira, D.; da Silva, R.P.; Maldonado, W., Jr.; Zerbato, C. Convolutional neural networks in predicting cotton yield from images of commercial fields. *Comput. Electron. Agric.* 2020, 171, 105307. [CrossRef]
- Nevavuori, P.; Narra, N.; Lipping, T. Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* 2019, 163, 104859. [CrossRef]

- 87. Chu, Z.; Yu, J. An end-to-end model for rice yield prediction using deep learning fusion. *Comput. Electron. Agric.* 2020, 174, 105471. [CrossRef]
- Nguyen, T.T.; Hoang, T.D.; Pham, M.T.; Vu, T.T.; Nguyen, T.H.; Huynh, Q.T.; Jo, J. Monitoring agriculture areas with satellite images and deep learning. *Appl. Soft Comput.* 2020, 95, 106565. [CrossRef]
- 89. Waldner, F.; Diakogiannis, F.I. Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sens. Environ.* 2020, 245, 111741. [CrossRef]
- 90. Wei, S.; Zhang, H.; Wang, C.; Wang, Y.; Xu, L. Multi-temporal SAR data large-scale crop mapping based on U-Net model. *Remote Sens.* 2019, 11, 68. [CrossRef]
- 91. Papadomanolaki, M.; Vakalopoulou, M.; Zagoruyko, S.; Karantzalos, K. Benchmarking deep learning frameworks for the classification of very high resolution satellite multispectral data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 2016, *3*, 83–88. [CrossRef]
- Sagan, V.; Maimaitijiang, M.; Bhadra, S.; Maimaitiyiming, M.; Brown, D.R.; Sidike, P.; Fritschi, F.B. Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. *ISPRS J. Photogramm.* 2021, 174, 265–281. [CrossRef]
- 93. Meng, S.Y.; Wang, X.Y.; Hu, X.; Luo, C.; Zhong, Y.F. Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery. *Comput. Electron. Agric.* 2021, 186, 106188. [CrossRef]
- 94. Xu, L.; Zhang, H.; Wang, C.; Zhang, B.; Liu, M. Crop classification based on temporal information using sentinel-1 SAR time-series data. *Remote Sens.* 2019, *11*, 53. [CrossRef]
- Gella, G.W.; Bijker, W.; Belgiu, M. Mapping crop types in complex farming areas using SAR imagery with dynamic time warping. ISPRS J. Photogramm. 2021, 175, 171–183. [CrossRef]
- Huang, Z.L.; Datcu, M.; Pan, Z.X.; Lei, B. Deep SAR-Net: Learning objects from signals. *ISPRS J. Photogramm.* 2020, 161, 179–193. [CrossRef]
- 97. Zhang, T.W.; Zhang, X.L.; Shi, J.; Wei, S.J. HyperLi-net: A hyper-light deep learning network for high-accurate and high-speed ship detection from synthetic aperture radar imagery. *ISPRS J. Photogramm.* **2020**, *167*, 123–153. [CrossRef]
- Zheng, Z.; Ma, A.; Zhang, L.P.; Zhong, Y.F. Deep multisensor learning for missing-modality all-weather mapping. *ISPRS J. Photogramm.* 2021, 174, 254–264. [CrossRef]
- 99. Ienco, D.; Interdonato, R.; Gaetano, R.; Minh, D.H.T. Combining sentinel-1 and sentinel-2 satellite image time series for land cover mapping via a multi-source deep learning architecture. *ISPRS J. Photogramm.* **2019**, *158*, 11–22. [CrossRef]
- Adrian, J.; Sagan, V.; Maimaitijiang, M. Sentinel SAR-optical fusion for crop type mapping using deep learning and google earth engine. *ISPRS J. Photogramm.* 2021, 175, 215–235. [CrossRef]
- Zhao, W.Z.; Qu, Y.; Chen, J.; Yuan, Z.L. Deeply synergistic optical and SAR time series for crop dynamic monitoring. *Remote Sens. Environ.* 2020, 247, 111952. [CrossRef]
- 102. Mahdianpari, M.; Salehi, B.; Rezaee, M.; Mohammadimanesh, F.; Zhang, Y. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens.* **2018**, *10*, 1119. [CrossRef]
- 103. Vali, A.; Comai, S.; Matteucci, M. Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sens.* **2020**, *12*, 2495. [CrossRef]
- 104. Kaya, A.; Keceli, A.S.; Catal, C.; Yalic, H.Y.; Temucin, H.; Tekinerdogan, B. Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* 2019, 158, 20–29. [CrossRef]
- Cai, E.; Baireddy, S.; Yang, C.; Crawford, M.; Delp, E.J. Deep transfer learning for plant center localization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 62–63.
- Espejo-Garcia, B.; Mylonas, N.; Athanasakos, L.; Fountas, S. Improving weeds identification with a repository of agricultural pre-trained deep neural networks. *Comput. Electron. Agric.* 2020, 175, 105593. [CrossRef]
- 107. Bosilj, P.; Aptoula, E.; Duckett, T.; Cielniak, G. Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. *J. Field Robot.* **2020**, *37*, 7–19. [CrossRef]
- 108. Abdalla, A.; Cen, H.; Wan, L.; Rashid, R.; Weng, H.Y.; Zhou, W.J.; He, Y. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Comput. Electron. Agric.* 2019, 167, 105091. [CrossRef]
- 109. Suh, H.K.; Ijsselmuiden, J.; Hofstee, J.W.; van Henten, E.J. Transfer learning for the classification of sugar beet and volunteer potato under field conditions. *Biosyst. Eng.* 2018, 174, 50–65. [CrossRef]
- 110. Fei-Fei, L.; Fergus, R.; Perona, P. One-shot learning of object categories. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 594–611. [CrossRef]
- Lake, B.M.; Salakhutdinov, R.; Tenenbaum, J.B. Human-level concept learning through probabilistic program induction. *Science* 2015, 350, 1332–1338. [CrossRef] [PubMed]
- 112. Chen, W.Y.; Liu, Y.C.; Kira, Z.; Wang, Y.C.F.; Huang, J.B. A closer look at few-shot classification. arXiv 2019, arXiv:1904.04232.
- 113. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In Proceedings of the International Conference on Machine Learning Deep Learning Workshop, Lille, France, 10–11 July 2015; Volume 2.
- 114. Li, Y.; Yang, J.C. Few-shot cotton pest recognition and terminal realization. Comput. Electron. Agric. 2020, 169, 105240. [CrossRef]
- 115. Hu, G.S.; Wu, H.Y.; Zhang, Y.; Wan, M.Z. A low shot learning method for tea leaf's disease identification. *Comput. Electron. Agric.* **2019**, *163*, 104852. [CrossRef]

- Liu, B.; Yu, X.C.; Yu, A.Z.; Zhang, P.Q.; Wan, G.; Wang, R.R. Deep few-shot learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 2018, 57, 2290–2304. [CrossRef]
- Sung, F.; Yang, Y.X.; Zhang, L.; Xiang, T.; Torr, P.H.; Hospedales, T.M. Learning to compare: Relation network for few-shot learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake, UT, USA, 18–23 June 2018; pp. 1199–1208.
- 118. Andrychowicz, M.; Denil, M.; Gomez, S.; Hoffman, M.W.; Pfau, D.; Schaul, T.; De Freitas, N. Learning to learn by gradient descent by gradient descent. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 3981–3989.
- 119. Gao, K.L.; Liu, B.; Yu, X.C.; Qin, J.C.; Zhang, P.Q.; Tan, X. Deep relation network for hyperspectral image few-shot classification. *Remote Sens.* **2020**, *12*, 923. [CrossRef]
- Haug, S.; Ostermann, J. A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014; pp. 105–116.
- 121. Lameski, P.; Zdravevski, E.; Trajkovik, V.; Kulakov, A. Weed detection dataset with RGB images taken under variable light conditions. In Proceedings of the International Conference on ICT Innovations, Skopje, Macedonia, 18–23 September 2017; Springer: Cham, Switzerland, 2017; pp. 112–119.
- 122. Wiesner-Hanks, T.; Stewart, E.L.; Kaczmar, N.; DeChant, C.; Wu, H.; Nelson, R.J.; Lipson, H.; Gore, M.A. Image set for deep learning: Field images of maize annotated with disease symptoms. *BMC Res. Notes* **2018**, *11*, 440. [CrossRef] [PubMed]
- 123. Wiesner-Hanks, T.; Wu, H.; Stewart, E.; DeChant, C.; Kaczmar, N.; Lipson, H.; Gore, M.A.; Nelson, R.J. Millimeter-level plant disease detection from aerial photographs via deep learning and crowdsourced data. *Front. Plant Sci.* 2019, 10, 1550. [CrossRef]
- 124. Liu, X.; Min, W.; Mei, S.; Wang, L.; Jiang, S. Plant disease recognition: A large-scale benchmark dataset and a visual region and loss reweighting approach. *IEEE Trans. Image Process.* **2021**, *30*, 2003–2015. [CrossRef]
- 125. Chiu, M.T.; Xu, X.; Wei, Y.; Huang, Z.; Schwing, A.G.; Brunner, R.; Khachatrian, H.; Karapetyan, H.; Dozier, I.; Rose, G.; et al. Agriculture-vision: A large aerial image database for agricultural pattern analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2828–2838.
- 126. Su, J.; Yi, D.; Su, B.; Mi, Z.; Liu, C.; Hu, X.; Xu, X.; Guo, L.; Chen, W.H. Aerial visual perception in smart farming: Field study of wheat yellow rust monitoring. *IEEE Trans. Ind. Inform.* 2021, 17, 2242–2249. [CrossRef]
- Abbas, A.; Jain, S.; Gour, M.; Vankudothu, S. Tomato plant disease detection using transfer learning with C-GAN synthetic images. *Comput. Electron. Agric.* 2021, 187, 106279. [CrossRef]
- 128. Khaki, S.; Pham, H.; Han, Y.; Kuhl, A.; Kent, W.; Wang, L. Deepcorn: A semi-supervised deep learning method for high-throughput image-based corn kernel counting and yield estimation. *Knowl. Based Syst.* **2021**, *218*, 106874. [CrossRef]
- 129. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Hartling, S.; Esposito, F.; Fritschi, F.B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* **2020**, 237, 111599. [CrossRef]
- Sa, I.; Chen, Z.; Popović, M.; Khanna, R.; Liebisch, F.; Nieto, J.; Siegwart, R. WeedNet: Dense semantic weed classification using multispectral images and MAV for smart farming. *IEEE Robot. Autom. Lett.* 2017, *3*, 588–595. [CrossRef]
- Sa, I.; Popović, M.; Khanna, R.; Chen, Z.; Lottes, P.; Liebisch, F.; Nieto, J.; Stachniss, C.; Walter, A.; Siegwart, R. WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming. *Remote Sens.* 2018, 10, 1423. [CrossRef]
- 132. Jiao, L.; Dong, S.; Zhang, S.; Xie, C.; Wang, H. AF-RCNN: An anchor-free convolutional neural network for multi-categories agricultural pest detection. *Comput. Electron. Agric.* 2020, 174, 105522. [CrossRef]