Review

Convolutional Neural Networks in Computer Vision for Grain Crop Phenotyping: A Review

Ya-Hong Wang and Wen-Hao Su * (ID)

College of Engineering, China Agricultural University, 17 Qinghua East Road, Haidian, Beijing 100083, China
* Correspondence: wenhao.su@cau.edu.cn

Abstract: Computer vision (CV) combined with a deep convolutional neural network (CNN) has emerged as a reliable analytical method to effectively characterize and quantify high-throughput phenotyping of different grain crops, including rice, wheat, corn, and soybean. In addition to the ability to rapidly obtain information on plant organs and abiotic stresses, and the ability to segment crops from weeds, such techniques have been used to detect pests and plant diseases and to identify grain varieties. The development of corresponding imaging systems to assess the phenotypic parameters, yield, and quality of crop plants will increase the confidence of stakeholders in grain crop cultivation, thereby bringing technical and economic benefits to advanced agriculture. Therefore, this paper provides a comprehensive review of CNNs in computer vision for grain crop phenotyping. It is meaningful to provide a review as a roadmap for future research in such a thriving research area. The CNN models (e.g., VGG, YOLO, and Faster R-CNN) used CV tasks including image classification, object detection, semantic segmentation, and instance segmentation, and the main results of recent studies on crop phenotype detection are discussed and summarized. Additionally, the challenges and future trends of the phenotyping techniques in grain crops are presented.

Keywords: grain crops; convolutional neural network; computer vision; phenotype detection

1. Introduction

Global food security remains an important issue for human development [1]. By 2050, the global population is likely to exceed 9 billion, which means that agricultural production will need to increase by at least 70% from its current level to meet the growing demand for food [2]. Grains are the main component of the human diet, and rice, wheat, corn, and soybean account for more than 80% of global grain production [3]. Intelligent perception of crop phenotypic information helps to achieve precise field management, such as the selection of new varieties of high-yield and high-quality crops, and the minimization of agricultural inputs without affecting crop output. Plant phenotypes are the recognizable morphological, physiological, and biochemical characteristics and traits resulting from gene-environment interactions, including plant structure, composition, growth, and development [4]. This means that phenotypic assessment not only involves the traits expressed by crop genes, but also reflects complex traits such as physiology, biochemistry, quality, stress resistance, or ones that are influenced by the external environment.

Computer vision (CV), when combined with pattern recognition algorithms and automatic classification tools, exhibits outstanding performance. Traditional plant phenotype detection relies on manual observation and measurement to obtain a description of the external morphology of the plant, and then assess the relationship between genes or external environment and phenotype. However, this approach can only detect individual traits from a small sample of crops, thus the acquisition process is inefficient and the amount of data available is very limited. With the increasing demand for high-volume plant phenotypic information, researchers urgently need high-precision, high-throughput, and low-cost techniques to replace traditional manual methods of obtaining relevant data.

A variety of imaging techniques are available to collect complex traits related to growth, yield, and adaptation to biotic or abiotic stresses (e.g., diseases, insects, water stress, and nutrient deficiencies), including color imaging (e.g., machine vision), imaging spectroscopy (e.g., multi-spectral and hyperspectral remote sensing), thermal infrared imaging, fluorescence imaging, 3D imaging, and laminar imaging [5].

Over the past few decades, computer vision has been widely applied to analyze the phenotypic characteristics of grain crops and thus ease the food supply problem. Although a review of the phenotypic assessment of grain crops based on computer vision was published in 2018, the research mainly summarized the application of traditional machine-learning algorithms such as the support vector machine (SVM) and the back-propagation neural network (BPNN) [6]. In addition, some researchers have reviewed the research on pest and disease analysis of crops [7], crop and weed identification [8], and physical and chemical phenotypic characteristics of crops [9], but they only mentioned a particular phenotyping task. Importantly, some new network architectures and strategies applied to the field of the convolutional neural network (CNN) and computer vision are rarely covered in the extensive reviews covering crop phenotype detection since2019. Several papers have been published in the last three years that provide comprehensive reviews of deep learning techniques for such computer vision tasks as image classification [10], object detection [11], and semantic and instance segmentation [12]. These reviews effectively summarize the basic principles, development history, and future trends for the latest CNNs in computer vision, but none of them provide information related to agriculture, which highlights a gap between these technological theories and phenotyping applications.

Focusing on the state-of-the-art CNN algorithms rather than traditional machine learning (the specific differences are shown in Figure 1), this study is an important early step in the search for phenotyping of grain crops. Given the importance of the four most productive grain crops (rice, wheat, maize, and soybean) in the world, the related work on computer vision-based CNN models for the detections of crop organs, crops in weeds, plant diseases, insect infestations, abiotic stresses, and grain varieties since 2019 has been reviewed. The goal is to provide a comprehensive overview of novel CNN models combined with CV for phenotype detection in grain crops and to provide researchers and breeders with clear guidance for related decisions. This will greatly boost the productivity of grain crops.
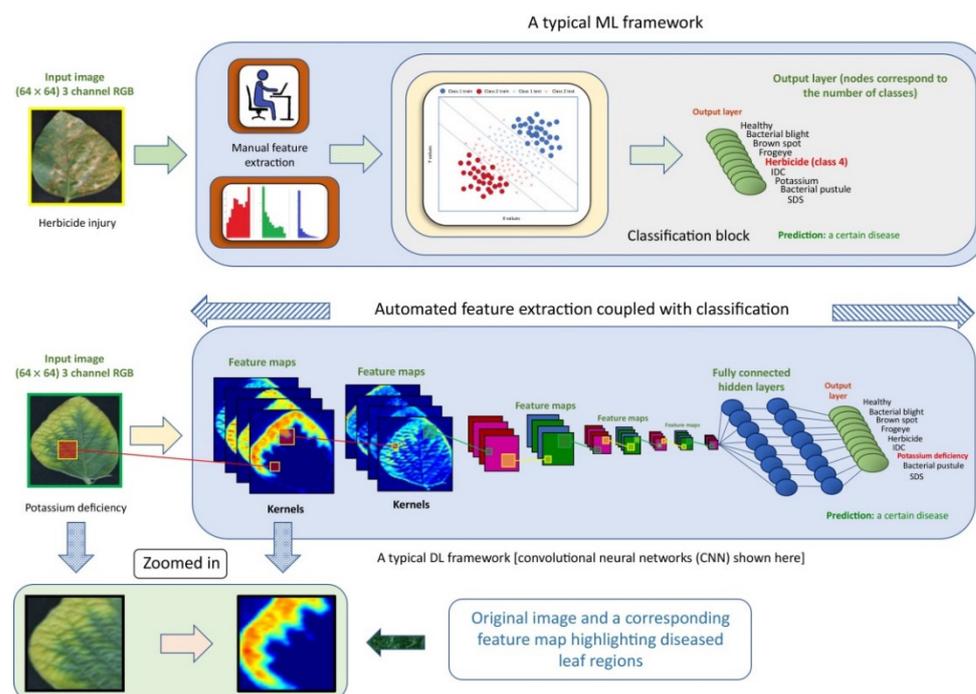

Figure 1. Key differences between machine learning (ML) and deep learning (DL) paradigms [13].

2. Computer Vision (CV) and Convolutional Neural Networks (CNNs)

2.1. CV

In recent years, both the hardware and software of CV systems have been significantly developed. The hardware, including cameras, lights, and communication devices, is the foundation of CV, while the software, such as image processing algorithms, is the core of the system. A typical image acquisition system is indispensable to illumination devices. The illumination devices can be divided into point light sources, strip light sources, ring light sources, backlight light sources, structure light sources, and combined light sources. These light sources can be further classified as light-emitting diode (LED) light sources, halogen light sources, and high-frequency fluorescent light sources. In addition, the camera can be characterized as a global shutter or a roll-up shutter camera.

2.2. CNN

Since 2012, CNNs have dominated solutions to CV tasks, showing superior performance over traditional machine-learning methods [14]. CNNs are deep learning architectures with spontaneous feature learning for image processing and image recognition. After the parameter optimization of training and learning, the CNN performs multiple layers of nonlinear transformations on the input data, continuously coupling the low-level features, and finally obtains a high-level semantic representation. Compared with traditional machine learning, a CNN can use a deeper neural network model to train the input data to simplify the data processing process.

A typical CNN consists of a convolutional layer, a pooling layer, and a fully connected layer [15]. The neurons in the convolutional layer are arranged in a matrix to form a multi-channel feature map. A neuron in each channel is connected to only a part of the feature map before that layer [16]. The final input of the neuron is obtained by convolving it with a convolution kernel and then using an activation function. CNNs emphasize weight sharing as a key component. Neurons located on the same channel feature map of the same convolutional layer are obtained by applying the same convolutional kernel to the previous feature map of the layer. Guided by local features in higher feature maps, the convolutional layer searches for links between them, while pooling layers combine data with the same semantics. Because the graphical information formed by adjacent positions may be slightly jittered, the pooling operation extracts the main information from the upper feature map. Maximum pooling and average pooling are common pooling operations. The model is able to keep translation and rotation invariant while preserving features [15]. After alternating between convolution and pooling, a fully connected layer often appears. Each neuron in the fully connected layer is connected to every neuron in the upper layer. All the information is combined to turn the multi-dimensional features into one-dimensional features, which are handed over to the final regressor and classifier to produce the final result.

2.3. CNNs Combined with CV Tasks

2.3.1. Image Classification

Image classification aims to assign predefined class labels to images. The CNN is currently the most popular neural network that combines a set of mathematical operations (e.g., convolution, pooling, and activation), using various connection schemes, such as plain stacking, start, and residual connections, to learn operational parameters from annotated images in order to classify image datasets (Figure 2). The current development of modern CNNs for image classification can be divided into three phases: (1) the appearance of modern CNNs (2012–2014); (2) the development and refinement of CNN architecture intensification (2014–2017); and (3) the introduction of reinforcement learning and artificial intelligence for CNN architecture design (start of 2017).

In 2012, the first modern CNN architecture named AlexNet was proposed. The algorithm demonstrated strong performance in image classification in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC 2012) competition in that year [17]. The report of this model introduces a new era of image classification and other CV tasks using

CNNs. From 2014 to 2017, researchers developed several representative CNNs such as the residual neural network (ResNet) [18], the visual geometry group network (VGG) [19], and the dense convolutional network (DenseNet) [20] for image classification. These CNNs significantly improved the learning ability and recognition complexity by using efficient computational algorithms and modified connectivity schemes. From 2017, more studies focused on the use of reinforcement learning to search for the best CNN architecture that could yield higher performance [21]. This process introduces a reinforcement learning framework to find the optimal convolutional image elements on small datasets, followed by stacking and transferring the resulting image elements in a different way to a large unknown dataset.

Researchers have investigated the mechanisms of CNNs for image classification. A recent study improved AlexNet to create a new variant (ZFNet) using a visualization tool. This tool is a framework integrated with CNNs that can map neuronal activity back to the input pixel space. Thus, pixel-level activations can be visualized after each convolutional layer, which is particularly useful for understanding the CNN mechanism for further upgrades. CNNs can learn general representations of images rather than features solely for classification. Subsequent research developed various gradient-based methods, including guided backpropagation, gradient-weighted class activation mapping (Grad-CAM), and layer-by-layer relevance propagation (LRP). Meanwhile, some general frameworks (e.g., LIME and occlusion maps) can also be used to display important image regions for classification results [22,23].

2.3.2. Object Detection

Object detection is defined as determining the location of objects in a given image and the class to which each object belongs. As shown in Figure 2, object detection using CNNs can be divided into two categories: single-level and two-level CNN architectures. In the early framework development, OverFeat is the most representative model [24], and won the localization task of the 2013 ILSVRC competition. Then, a series of region-based region-convolution neural network (R-CNN) frameworks was introduced, including the original R-CNN [25], Fast R-CNN [26], and Faster R-CNN [27]. There are three key techniques in the RCNN architectures, including the region proposal network (RPN), region of interest (ROI) pooling operation, and multi-task loss function. The R-CNN family has been widely adopted as object detectors for various domain datasets.

2.3.3. Semantic and Instance Segmentations

Semantic segmentation aims to assign a class to each pixel in an image, but objects in the same classes are not distinguished. Instance segmentation outputs the mask and class of the target. Typically, CNN architectures for semantic and instance segmentations can be divided into two categories, including encoder-decoder-based frameworks and detection-based frameworks, as shown in Figure 2. The encoder-decoder-based model is the most primitive intelligent image segmentation network for improving segmentation accuracy. In the encoder stage, the CNN extracts semantic features from input samples. In the decoder stage, deconvolution is used to assign the extracted features to the label of each pixel. Representative models based on encoder-decoder include full convolutional networks (FCNs) [28], DeepLab [29], and U-Net [30]. Frameworks including R-CNN, Faster R-CNN, and Mask R-CNN have been widely used for instance segmentation [31,32].
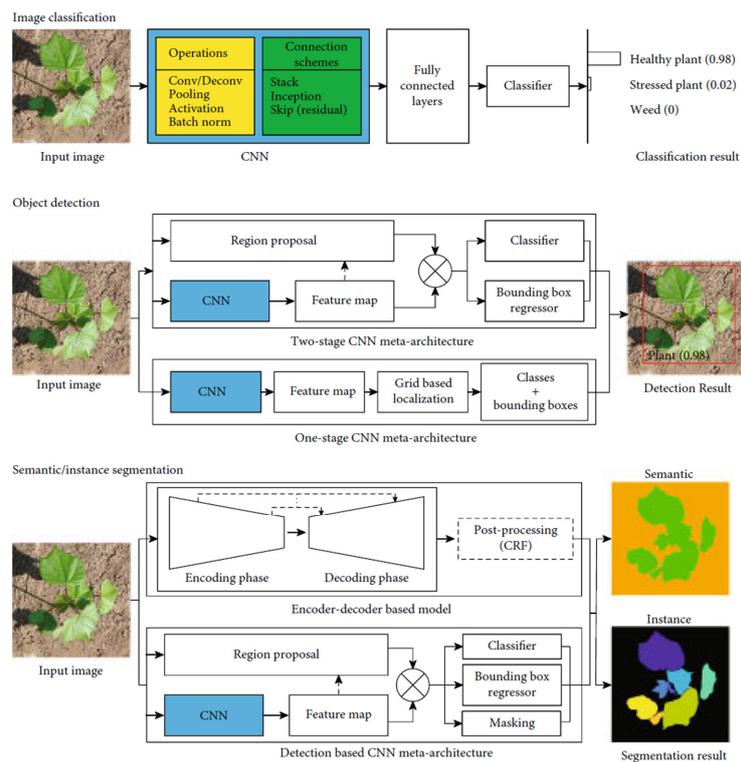
Figure 2. Diagrams of CNN architecture mechanisms for image classification, object detection, and semantic and instance segmentation [33].

3. Advances in Phenotyping of Four Grain Crops Based on CV and CNN

In grain crops, conventional inbreeding and artificial breeding based on molecular and genomic engineering are closely dependent on phenotypic information, which remains a bottleneck limiting crop breeding [34,35]. From field to table, grain crop phenotypes play an important role in enhancing crop germplasm, strengthening breeding, and evaluating commercial performance. Related researchers have invested a lot of effort in developing high-throughput and low-cost advanced phenotyping techniques. Particularly widely acknowledged is the development of CNNs combined with CV technology, which marks a new stage in crop phenotype detection. One of the challenges in breeding grain crops is to improve yield potential and quality stability [36]. However, traditional phenotyping methods based on manual measurements are typically labor-intensive and time-consuming when assessing multiple traits of crops [37]. The combination of cutting-edge CNNs and CV technology can achieve high-throughput screening of high-quality crop varieties, accurate yield prediction, automatic field weed detection, and early automatic diagnosis of pests and diseases, all of which are essential for the study of crop yield and quality enhancement.

3.1. Crop Organ Detection and Counting

Recent advances in CV and breakthroughs in deep learning have created new opportunities for the detection and counting of crop organs [38]. Traditional crop organ phenotypic information was obtained by manual measurement, such as measuring crop height and leaf width with a straightedge or counting using the naked eye. This is not only time-consuming and labor-intensive, but also has a limited variety of extracted features and low precision. CNN-based methods have shown promising results compared to traditional methods for crop selection in breeding programs [39]. Three methods, including object detection, semantic segmentation, and instance segmentation, are proposed for recognizing and counting the organs of the four major grain crops.

An object detection method which integrates the feature pyramid network (FPN) into the Faster R-CNN network has been successfully used for counting rice spikes [40]. Li et al. [41]

investigated the performance of Faster R-CNN and RetinaNet in predicting the number of wheat spikes at different growth stages. The RetinaNet model achieved higher accuracy for wheat spikes at the filling and maturity stages. Compared to Faster R-CNN and RetinaNet, Cascade R-CNN obtained a higher average precision (AP) of 89.6 for the detection and counting of soybean flowers and seeds [42]. You only look once (YOLO)v4 architecture was used to improve the detection speed and accuracy of wheat spikes [43]. TasselNet (ResNet34) was then established to detect the tassels of maize at different stages [44]. The backbone part of YOLOv4 was enhanced by adding a dual spatial pyramid pool (SPP) network to boost feature learning and broaden the perceptual domain of the convolutional network. The results obtained showed the superiority of the detection module compared to the early methods using SVM classifiers (Lu et al. [45], 2015) and neural network intensity model (Lu et al. [46], 2016).

Many studies investigated semantic segmentation based crop organ detection and counting. Sadeghi-Tehran et al. [47] not only successfully identified and quantified the number of wheat spikes in RGB images taken under initial natural field conditions, but also developed an efficient CV and CNN system based on DeepCount. The method used simple linear iterative clustering (SLIC) to segment images into super pixels and constructed a reasonable feature model for the semantic segmentation of wheat spikes. The results indicated that the model was able to detect the total number of wheat spikes in an image and estimate the number of spikes per square meter with a maximum accuracy of 98%. In another study, Xiong et al. [48] proposed a simple and effective contextual extension of TasselNet-TasselNetv2-that could significantly improve the performance of local regression networks. Experiments showed that TasselNetv2 was faster than TasselNet. Meanwhile, the classical model of semantic segmentation based on CNNs was used to detect the corn cob (Kienbaum et al. [49]). The Mask R-CNN model was used to extract shape parameters including asymmetry, ellipticity, and length of cobs, achieving an accuracy of about 100% for maize cob phenotypic parameters. It was found that the number of kernels in a corn cob image can be accurately estimated by DeepCorn (Khaki et al. [50]). In their work, DeepCorn uses VGG-16 as the backbone for feature extraction to merge elemental maps from multiple scales of the network, making it robust to image scale variations. DeepCorn successfully counted the kernels on a cob, regardless of their orientation and illumination conditions. In addition to this, Yang et al. [51] proposed a novel synthetic image generation and enhancement method based on domain randomization. The study used the Mask R-CNN model combined with transfer learning to perform the semantic segmentation of soybean seed images and successfully obtained specific organ phenotype parameters, which further deepens the application of CNNs in semantic segmentation tasks.

A noticeable concern is that although CNNs could provide accurate semantic masks, the counting accuracy can still suffer from inaccurate postprocesses. To address this concern, studies explored the use of instance segmentation CNNs that can directly segment individual objects in images [52]. For instance, a sophisticated soybean phenotypic measurement algorithm, named soybean phenotypic measurement instance segmentation (SPM-IS), was developed, enabling more rapid and accurate acquisition of phenotypic data for soybean stems, pods, and seeds (Li et al. [53]) This study used the Resnet-101-FPN model and SPM-IS algorithms to perform instance segmentation on images to measure the length and width of target objects to extract soybean phenotypic data. The test results showed that the mask MAP of pods, stems, and seeds were 95.7%, 93.5%, and 94.6%, respectively.

Faster R-CNN, Mask R-CNN, RetinaNet, and VGG have been widely studied with regard to the detection and counting of organs of grain crops, as shown in Table 1. Some strategies such as SPP, SLIC, and domain randomization have been added to the model training for the first time to achieve feature enhancement. In general, object detection is more widely used than semantic and instance segmentation in this area, but image classification was used less often in crop organ identification and counting. It is encouraging that some counting methods based on 3D image sequences or videos are emerging to provide a strategy to solve the above problems [54].

Table 1. Summary of major CNN-combined-with-CV tasks for crop organs images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Object detection	Rice	Counting of grain per panicle	RGB	Faster R-CNN with FPN	796	99.4%	Deng et al. [40]
	Wheat	Ear recognition	RGB	RetinaNet with FPN	52,920	98.6%	Li et al. [41]
	Wheat	Detection of head	RGB	YOLOv4 with dual SPP	3432	94.5%	Gong et al. [43]
	Maize	Tassel counting	RGB	TasselNet (ResNet34)	361	88.97%	Zou et al. [44]
	Soybean	Flower and seedpod detection	RGB	Cascade R-CNN, RetinaNet, Faster R-CNN.	76,524	AP = 89.6% AP = 83.3% AP = 88.7%	Pratama et al. [42]
	Soybean	Seed counting	RGB	TCNN	32,126	MAE = 13.21 MSE = 17.62	Li et al. [55]
	Soybean	Counting of seeds	RGB	VGG-Two	37,563	MAE = 0.6 MSE = 0.6	Ying et al. [56]
Semantic segmentation	Wheat	Counting of spikes	RGB	TasseINetv2	675,322	91.01%	Xiong et al. [48]
	Wheat	Quantification of spikes	RGB	VGG	580,000	98%	Sadeghi-Tehran et al. [47]
	Maize	Corn kernel counting	RGB	VGG-16	19,848	90.48%	Khaki et al. [50]
	Maize	Analysis of cob geometry	RGB	Mask R-CNN	19,867	100% for length 99% for diameter	Kienbaum et al. [49]
Instance segmentation	Soybean	Phenotype measurement	RGB	ResNet-101 with FPN	3207	MAP = 95.7%	Li et al. [53]

RGB—red, green, blue; CNN—convolution neural network; R-CNN—region-convolution neural network; FPN—feature pyramid networks, YOLO—you only look once; ResNet—residual neural network; VGG—visual geometry group network; TCNN—two-column convolution neural network; SPP—spatial pyramid pooling; AP—average precision; MAE—mean absolute error; MSE—mean squared error; MAP—mean average precision.

3.2. Weed and Crop Recognition and Segmentation

Weeds in a field compete with crops for nutrients, sunlight, and growing space. They need to be removed in time to avoid affecting crop yields [57]. Early applications of machine-learning methods to solve weed recognition problems generally used the color co-occurrence matrix (CCM) to extract features in terms of hue, color saturation, and intensity, or morphological and color features as input to the classifier [58]. However, the leaves of different plants often have the same color and shape, and it is challenging to identify weeds by the difference of leaf features. Traditional methods select artificially designed features to be extracted for distinction, which only performs well on specific datasets. With advances in intelligent sensing technology, the CNN-combined-with-CV technique has emerged as a promising tool for accurate and real-time detection of weeds and crops in the field

Rice is a crop with fixed rice row spacing, which can be identified by the location identification method. Lin et al. [59] developed a Faster R-CNN model to determine the specific row spacing parameters and successfully detected rice seedlings from weeds with an accuracy of 89.8%. Wang et al. [60] proposed a new method for the recognition of rice seedling rows based on row vector grid classification. In their research, seedling feature extraction and row vector grid classification were built into an end-to-end CNN model. The method successfully realized crop recognition in complex weed scenarios. The Faster R-CNN model was also used to distinguish between weeds and maize [61]. This study proposed an architecture using a VGG19 pre-trained network for distinguishing maize seedlings from weeds under complex field conditions. The results revealed that Faster R-CNN model has great potential for plant detection. Additionally, Jiang et al. [62] proposed a graph convolutional network (GCN) recognition method based on a similar approach. The GCN graph was constructed using the extracted CNN features of weeds and their Euclidean distances for maize and weed recognition. The results show that the GCN-ResNet-101 method achieved an accuracy score of 97.80%, which was better than state-of-the-art methods including AlexNet, VGG16, and ResNet-101.

Semantic segmentation has also been applied to the management of weeds. In a recent study, a multi-task semantic segmentation-convolutional neural network (MTS-CNN) model was designed for detecting crops and weeds using one-stage training [63]. This approach has heightened the correlations between the crop and weed classes, so that the object (crop and weed) region is trained intensively with the highest segmentation accuracy. Weirong et al. [64] proposed an improved Mask R-CNN-based algorithm for maize seedling segmentation. The model was trained using ResNeXt50/101-FPN as a feature extraction network. The average recognition accuracy of the model was higher than 94.7%. Furthermore, Zhang et al. [65] developed a weed classification model based on the YOLOV3-tiny network. In the study, a real-time detection system for field weeds based on unmanned aerial vehicles (UAVs) and mobile devices was designed to detect five kinds of weeds. Furthermore, Haq [66] and Babu and Ram [67] have conducted taxonomic studies on grasses and broadleaf weeds in soybean. The network architectures used were CNNs with learning vector quantization (LVQ), and a deep residual convolutional neural network (DRCNN). Both methods achieved over 97% accuracy for the individual targeting of two weed species.

In summary, the discrimination of crops such as rice and maize from complex weeds depends on the correct identification and localization of the plant by the model. Researchers have proposed many CNN-based solutions, most of which are implemented using object detection and semantic segmentation (the related studies are tabulated in Table 2). All of these results far exceed the accuracy achieved by a wide range of methods with artificially designed features. In recent years, researchers have achieved high recognition and segmentation accuracies on rice and maize image datasets by using classical networks such as ResNet and Faster R-CNN, or by building other shallow networks. Other studies were carried out on the classification of weed species based on supervised and semi-supervised learning methods [68,69]. In the future, advanced network models and more comprehensive datasets are needed to enable the identification of multiple crops and common weeds [70].

Table 2. Summary of major CNN-combined-with-CV tasks for weed and crop images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Image classification	Wheat	Identification of weed species	RGB	YOLOv3-tiny	2000	mAP = 72.5%	Zhang et al. [65]
	Soybean	Weed detection	RGB	CNN-LVQ	15,000	99.44%	Haq [66]
	Soybean	Weed classification	RGB	DRCNN	15,336	97.25%	Babu and Ram [67]
Object detection	Rice	Seedling rows recognition	RGB	ResNet-18	4500	88.54%	Wang et al. [60]
	Rice	Location of seedlings	RGB	Faster R-CNN	240	89.8%	Lin et al. [59]
	Maize	Seedling detection	RGB	Faster R-CNN (VGG19)	32,354	97.71%	Quan et al. [61]
	Maize	Weed recognition	RGB	GCN-ResNet-101	6000	97.8%	Jiang et al. [62]
	Maize	Plant detection	RGB	Faster R-CNN	211	99.8% at 0.5 Intersection over Union	Zhang et al. [71]
Semantic segmentation	Rice	Segmentation of weeds	RGB	MTS-CNN	224	96.48%	Kim and Park [63]
	Maize	Seedling and core detection	RGB	Mask R-CNN (ResNet50/101-FPN)	1800	94.7%	Weirong et al. [64]

GCN—graph convolutional network; CNN-LVQ—convolutional neural network-learning vector quantization; DRCNN—deep residual convolutional neural network; MTS-CNN—multi-task semantic segmentation-convolutional neural network.

3.3. Crop Disease Detection and Classification

The intelligent detection of plant diseases has received increasing attention in recent years. Crop diseases negatively affect agricultural production [72]. Early detection and control of crop diseases play a crucial role in the management of and decision making involved in agricultural production. Traditional machine learning approaches to feature analysis of crop photographs can detect diseases earlier than human observation. Nevertheless, such methods focusing on a limited number of crops were usually performed on small data sets. In recent years, methods based on deep learning and image technologies have been widely used in plant pathology.

CNNs have been successfully used for the classification and detection of crop diseases. Sharma et al. [73] developed a CNN model based on transfer learning to classify diseases in rice leaf images. Based on the disease features, Krishnamoorthy et al. [74] successfully distinguished three invasive rice diseases, including leaf blast, white leaf blight, and brown spot, from healthy rice leaves, with an accuracy of 95.67%. Singh and Arora [75] and Kumar and Kukreja [76] developed seven CNN models to classify wheat diseases including powdery mildew, stem rust, and leaf rust. Compared with VGG16, VGG19, AlexNet, ResNet-34, ResNet-50, and ResNet-18, ResNet101 achieved the highest accuracy of 98.6%. Similarly, Jiang et al. [77] adopted the PlantVillage dataset to pretrain several CNN models based on transfer learning. Figure 3 shows the comparison of CNNs used in this study in terms of accuracy, memory, and processing time, which will provide a reference direction for other disease diagnosis.
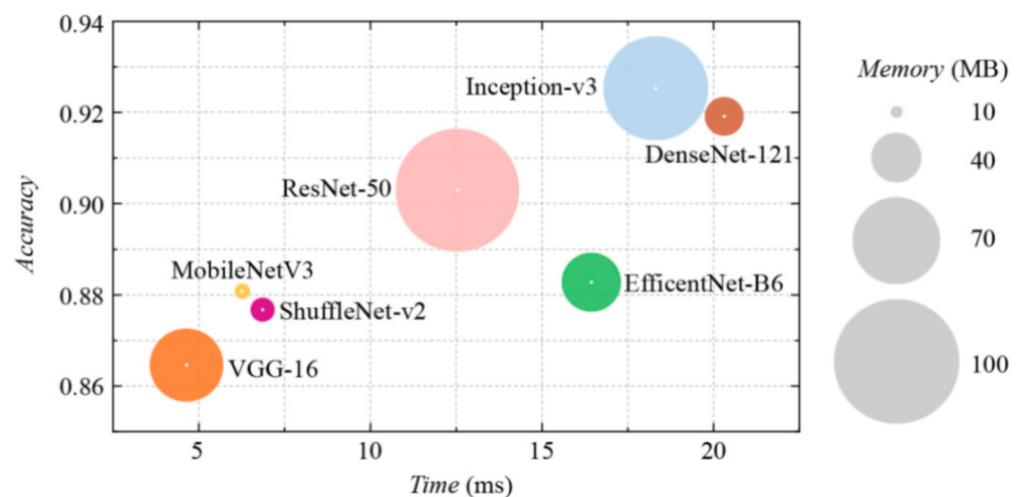

Figure 3. Performance of different models for disease diagnosis [77].

Object detection combined with the classical Faster R-CNN model plays a great role in the detection of grain crop diseases. Bari et al. [78] found a solution for the real-time detection of rice leaf diseases using Faster R-CNN to precisely localize the target. Results showed that the approach has accuracies of 98.09%, 98.85%, and 99.17% for automatic detection of three rice blast, brown spot, and hispa, respectively. In another study, Zhou et al. [79] proposed a fast rice disease detection method based on the K-mean clustering algorithm (FCM-KM) and fast R-CNN. FCM-KM was optimized using the chaos-based dynamic population firefly algorithm and maximum minimum distance. Zhang et al. [80] designed a multi-feature fusion faster R-CNN (MF3R-CNN) model for the detection of soybean leaf disease, with an average accuracy of 83.34%. Compared to the studies of Shrivastava et al. [81] (2017) and Pires et al. [82] (2016) using the k-nearest neural network and local descriptors to distinguish the diseased from the healthy, the model has practical implications for multiple disease identifications.

Grain crop diseases in complex environments have been successfully detected based on semantic segmentation. For instance, Ennadifi et al. [83] conducted a mask R-CNN to segment wheat spikes from the background. Then, a DenseNet121 model combined

with gradient-weighted class activation mapping (GradCAM) was used for localizing the diseased areas on wheat spikes in an unsupervised manner, yielding an accuracy of 93.47%. Nevertheless, wheat disease classification is susceptible to various visual disturbances. Lin et al. [84] proposed an M-bCNN model for the classification of wheat leaf diseases, achieving a test accuracy of 90.1%. Su et al. [85] developed a Mask-RCNN model to evaluate Fusarium head blight (FHB) severity with an accuracy of 77.19% (Figure 4). In their study, a ResNet-101 network-based FPN was used as the backbone of Mask-RCNN to segment wheat spikes and diseased areas, yielding accuracies of 77.76% and 98.81%. On this basis, the FPN based on the ResNet network was further upgraded as the backbone of BlendMask networks for the severity assessment of wheat FHB [86]. The newly constructed model demonstrated outstanding performance in the identification of wheat spikes occluded by awns, which is more concise and efficient than the Mask R-CNN.
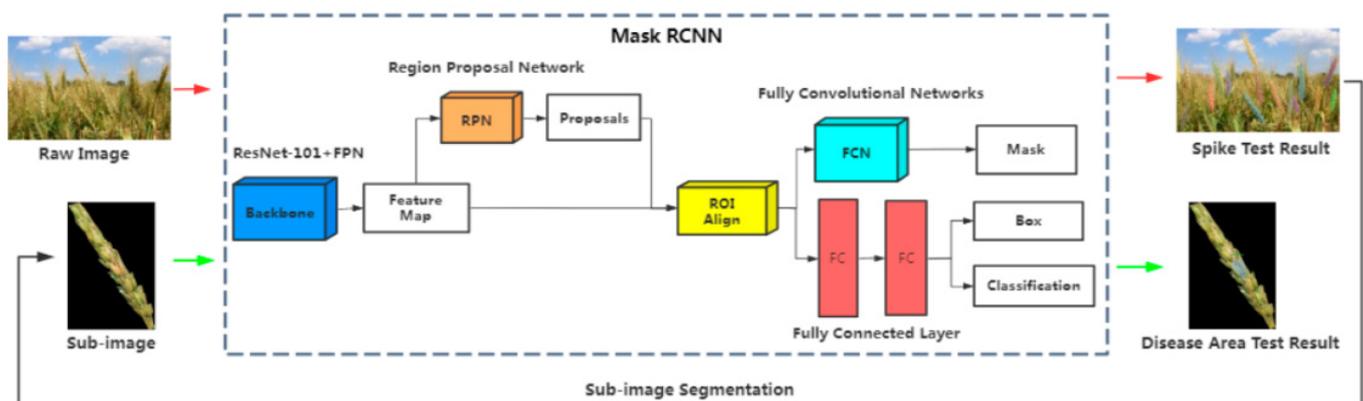

Figure 4. The architecture of the mask region convolutional neural network (Mask-RCNN) approach for wheat Fusarium head blight (FHB) disease assessment [85].

In general, image classification has more comprehensive applications than object detection and segmentation tasks in crop disease detection. The Mask R-CNN model not only allows for semantic segmentation, but also allows for more efficient analysis of disease severity levels. Faster R-CNN, when used as a tool for object detection, focuses more on spot location identification. When combined with FCM-KM, the results obtained are more comprehensive and convenient. Detailed comparison results are shown in Table 3. The CNN model combined with FCM-KM, GradCAM, and other strategies have been groundbreakingly optimized to provide new idea for crop disease detection. In addition, it has been found that a transfer-learning method employing retuning of all parameters produced the highest accuracy [77].

Table 3. Summary of major CNN-combined-with-CV tasks for crop disease images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Image classification	Rice	Hispa disease classification	RGB	CNN	1000	94%	Sharma et al. [73]
	Rice	Leaf disease recognition	RGB	InceptionResNetV2	5200	95.67%	Krishnamoorthy et al. [74]
	Wheat	Classification of powdery mildew disease	RGB	N-CNN	450	89.9%	Kumar and Kukreja [76]
	Wheat	Detection of healthy and unhealthy wheat	RGB	ResNet101, VGG-19, AlexNet	750	98.6% 96.6% 92.6%	Singh and Arora [75]
	Maize	Leaf disease identification	RGB	GoogleNet	8604	98.55%	Krishnamoorthi et al. [87]
	Maize	Detection of leaf diseases	RGB	MAF-ResNet50	59,778	97.41%	Zhang et al. [88]
	Maize	Disease classification	RGB	CNN with BiLSTM	29,065	99.02%	Hasan et al. [89]
	Maize	Classification of leaf diseases	RGB	DeepForest (gcForest)	400	96.25%	Arora et al. [90]
	Soybean	Leaf disease classification	RGB	SoyNet	17,240	98.14%	Karlekar and Seal [91]
Object detection	Rice	Leaf disease detection	RGB	Faster R-CNN with RPN	16,800	98.09% for blast, 98.85% for brown spot, 99.17% for hispa.	Bari et al. [78]
	Rice	Disease detection	RGB	Faster R-CNN with FCM-KM	3010	96.71% for blast, 97.53% for bacterial blight, 98.3% for blight.	Zhou et al. [79]
	Soybean	Leaf disease detection	RGB	Muti-feature fusion Faster R-CNN	2230	96.43% for virus, 87.76% for frogeye spot, 65.63% for bacterial spot.	Zhang et al. [80]

Table 3. *Cont.*

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Semantic segmentation	Wheat	Classification of leaf diseases	RGB	M-bCNN	83,260	96.5%	Lin et al. [84]
	Wheat	Disease classification and localization	RGB	Mask R-CNN (DensNet12)	1163	93.47%	Ennadifi et al. [83]
	Wheat	Ear disease identification	RGB	SimpleNet	1205	93% for blotch 93% for scab	Bao et al. [92]
	Wheat	Yellow rust disease recognition	RGB	PSP Net	5580	98%	Pan et al. [93]
	Maize	Detection of gray leaf spot severity levels	RGB	CNN	1500	95.33%	Baliyan et al. [94]
Instance segmentation	Wheat	Evaluation of resistance to fusarium head blight (FHB)	RGB	Mask-RCNN (ResNet-101, FPN, RPN)	17,340	77.76% for spike, 98.81% for diseased area, 77.19% for FHB severity.	Su et al. [85]
	Wheat	Severity assessment of FHB	RGB	Tandem Dual BlendMask (ResNet-50, FPN)	3754	85.56% for spike, 99.32% for diseased area, 91.80% for FHB severity.	Gao, Wang, Li and Su [86]

N-CNN—neonatal convolutional neural network; MAF—multi-pathway activation function; M-bCNN—matrix-based convolutional neural network; RPN—region proposal network; PSP Net—pyramid scene parsing network; BiLSTM—bi-directional long short-term memory.

3.4. Crop Insect Infestation Detection

Pests have a significant impact in crop destruction [95]. However, the extensive use of chemicals such as pesticides to control pests has had adverse effects on agro-ecosystems [96]. Traditional methods use chlorophyll histograms to detect discoloration caused by pests, or SVM combined with special algorithms to identify the presence of pests [97,98]. For these methods, segmentation becomes difficult if the background contains distractions such as other leaves and plants. In addition, designing artificial features such as color histograms and texture features requires expertise, which is difficult to apply universally. Lately, numerous CNN-based pest identification methodologies have been presented in the computer vision field, which have showed brilliant execution in early pest control.

For insect infestation classification tasks, many cutting-edge models and strategies have been continuously developed in recent years, which in turn resulted in the proposal of more efficient deep networks. For example, four different CNN models, including VGG16, VGG19, InceptionV3, and MobileNetV2, were applied to the detection of maize leaves infected by fall armyworms (faw) [99]. The study found that InceptionV3 and MobileNetV2 performed better than the other models, with identification accuracies of 100%. Moreover, Tetila et al. [100] and Abade et al. [101] innovatively used the simple linear iterative clustering (SLIC) strategy and the NemaNet model, respectively, to classify pest-infested soybean images and both showed extremely promising results

Object detection is a computer vision task that involves the identification of an object class with its location in the image. On this basis, Li et al. [102] developed a Resnet-50 with the region proposal network (RPN) for pest identification in wheat fields, achieving the accuracies of 90.88%, 88.76%, and 70.2% for wheat sawfly, wheat aphid, and wheat mite, respectively. In addition, the Faster R-CNN model was effectively applied to detect pest infection in grain crops [103]. Furthermore, Verma et al. [104] developed three popular CNN models to identify pests in soybean. YOLO v5 exhibited better performance than YOLOv3 and YOLOv4 in pest detection and recognition.

In summary, CNN-based network models including VGG, Faster R-CNN, and YOLO were effectively used for the detection of crop insect infestations. The VGG model mainly focuses on the species of the insect infestations, while Faster R-CNN and YOLO models are more often used for the identification and localization of the sites of infection, as shown in Table 4. The development of models using SLIC and models with RPN, or models that are currently popular but not used in this field, is a new direction for solving pest problems.

Table 4. Summary of major CNN-combined-with-CV tasks for crop insect infestations images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Image classification	Maize	Identification of leaves infected by fall armyworms	RGB	InceptionV3 MbileNetV2	11,280	100% 100%	Ishengoma et al. [99]
	Soybean	Classifying and counting of insect pests	RGB	DensNet-201 ResNet-50 Incetion-Resnet-v2	10,000	94.89% 93.78% 93.40%	Tetila et al. [100]
	Soybean	Identification of nematodes	RGB	NemaNet	3063	96.76% for FS 98.82% for TL	Abade et al. [101]
Object detection	Wheat	Pest localization	RGB	Resnet-50 with RPN	519,752	83.23%	Li et al. [102]
	Wheat	Recognition and counting of mites	RGB	ZFnet-5 VGG-16	546	94.6% 96.4%	Chen et al. [105]
	Maize	Pest detection	RGB	Faster R-CNN with RPN	15,000	97.53%	Sheema et al. [103]
	Soybean	Insect identification	RGB	YOLO v3, v4, and v5	3710	99.5% for the best AP	Verma et al. [104]

FS—full-scale; TL—time limit.

3.5. Abiotic Crop Stress Phenotype Assessment

Abiotic stresses, such as nutrient deficiency, drought, temperature, and salinity stresses, are major challenges for agriculture, and they lead to a significant reduction in crop growth and productivity [106]. The stress phenotyping assessment is an important tool for improving crop stress resistance, which can be divided into four stages: (1) identification (presence of stress); (2) classification (type of stress); (3) quantification (severity of stress); and (4) prediction (likelihood of stress occurrence) [13]. Although traditional machine-learning methods such as SVM, artificial neural networks (ANN), and Random Forests are often used to study abiotic stress phenotypes of crops [107], the development of deep CNN offers new opportunities to advance in this field. Image classification combined with CNN can be effectively used for abiotic crop stress detection. For rice crops, Nitrogen (N) concentration is a key indicator of health status. Sethy et al. [108] proposed a CNN-based method for predicting N deficiency stress in rice. They used six leading CNN architectures including ResNet-18, ResNet-50, GoogleNet, AlexNet, VGG-16, and VGG-19 to predict nitrogen deficiency. ResNet-50 +SVM outperformed the other five CNN-based classification models with an accuracy of 99.84%. Additionally, Wang et al. [109] and Rizal et al. [110] developed the Densenet-121 model and the ResNet-50 model respectively to evaluate the nutrient deficiencies of the rice leaves affected by three different types of nutrient deficiencies including N, phosphorus (P), and potassium (K), with an accuracy of over 97%. Furthermore, water stress affects the normal growth of grain. Zhuang et al. [111] developed a multi-scale CNN architecture with 2-Covs Units for the assessment of the water stress severity of maize, which realized automatic detection and severity quantification of water stress through computer vision techniques in a non-destructive way.

To sum up, in the current research results as shown in Table 5, image classification is more widely used for abiotic stress assessment than other computer vision tasks. Advanced CNN models such as VGG, YOLO, and Mask R-CNN have not been developed for this research area. It is worthwhile to expect that imaging modalities (e.g., hyperspectral imaging) combined with CNN will provide a new idea for phenotypic stress detection [112].

Table 5. Summary of major CNN-combined-with-CV tasks for abiotic crop stress images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Image classification	Rice	Prediction of nitrogen deficiency	RGB	ResNet-50 +SVM	5790	99.84%	Sethy et al. [108]
	Rice	Classification of nutrient deficiency	RGB	Densenet-121	1500	97%	Wang et al. [109]
	Rice	Nutrient deficiency evaluation	RGB	ResNet-50	1156	98%	Rizal et al. [110]
	Maize	Water stress recognition	RGB	CNN +SVM	18,040	88.41%	Zhuang et al. [111]

SVM—support vector machines.

3.6. Crop Seed Variety Classification

As a key input for crop production, seed is of great economic value, and its varietal classification is crucial for maintaining crop yield and varietal purity [113]. However, the phenotypic characteristics of different varieties of grain crop seeds are very similar, with significant overlap in morphology and color. Traditional seed variety classification usually requires manual annotation and judgment by experts in the agricultural field, which is very inefficient. Therefore, it is necessary to explore reliable methods to improve classification efficiency.

In the current research results, crop seed varieties can be classified effectively based on CNN technology. For example, Laabassi et al. [114] utilized five standard CNN structures (such as DensNet201, Inception V3, and MobileNet) trained based on transfer learning to classify wheat seeds into four varieties (Simeto, Vitron, ARZ, and HD), with the best classification accuracy of 95.68% attributed to DensNet201 architecture. Javanmardi et al. [115]

successfully classified nine corn seed varieties based on a VGG-16 pre-trained CNN model. Gao et al. [116] proposed a CNN based variety classification model for multiple growth periods of wheat. In the study, the CMPNet achieved high classification precision at the seed stage of wheat (Specific performance index shown in Figure 5) based on ResNet and SENet. In addition, Velesaca et al. [117] utilized a Mask R-CNN architecture to perform instance segmentation of maize seed samples. Meanwhile, some other typical models were then developed in the study, which showed that the CK-CNN achieved the best robustness and stability compared with VGG16 and ResNet50.
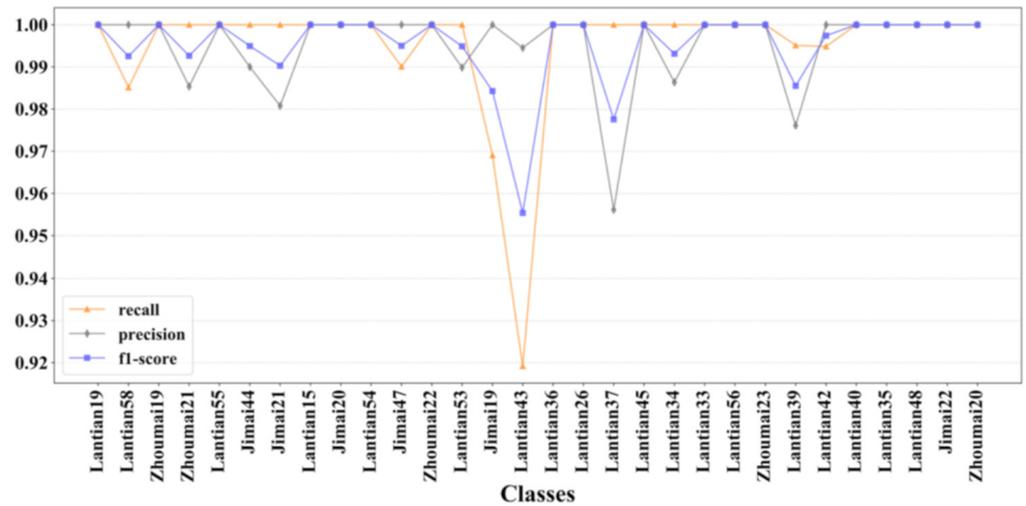

Figure 5. Evaluation index of CMPNet model classification [116].

It is worth mentioning that the newly improved model architectures combined with transfer learning such as P-ResNet showed the best accuracy to classify maize seeds in a non-destructive, fast and efficient manner [118]. The process is given in Figure 6. The result highlighted the advantages of transfer learning and its potential in deep learning, providing new solutions for CNN-based computer vision and spectroscopic techniques for seed classification and detection.
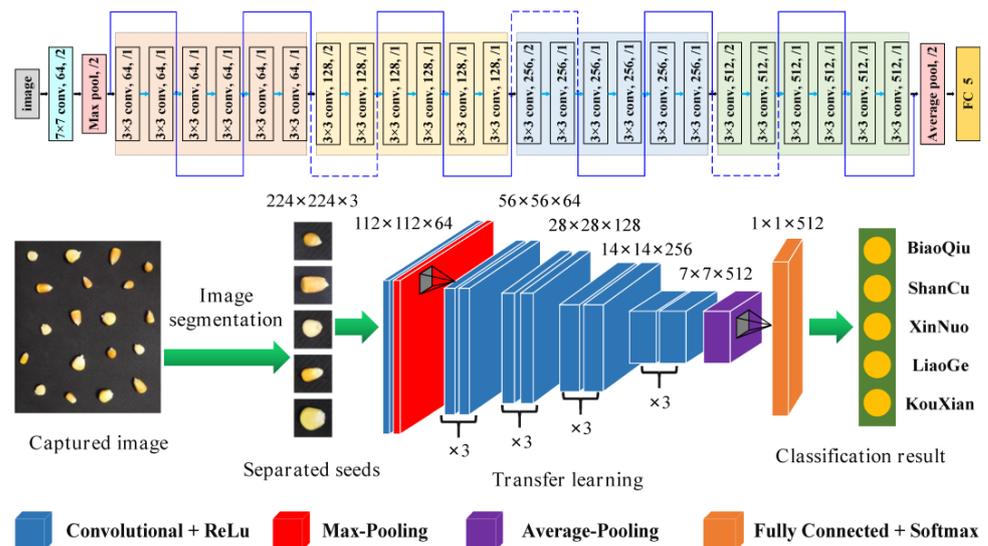

Figure 6. Process of transfer learning and classification of maize seeds [118].

In conclusion, various CNN models all have advantages and disadvantages for classifying crop seed varieties (the results of comparison are tabulated in Table 6). The classic DensNet and the novel corn kernel-CNN (CK-CNN) have higher accuracy than other models. Furthermore, the proposed methods, such as transfer learning and gradient-weighted

class activation mapping (Grad-CAM) techniques, provide new perspectives to maintain the classification accuracy and robustness of the model. In addition to computer vision, thermal imaging [119] and hyperspectral detection techniques [120] have also achieved great success in variety identification based on CNNs over the last decade.

Table 6. Summary of major CNN-combined-with-CV tasks for crop seeds images.

Vision Task	Crop	Phenotyping Task	Image Type	Model	Number of Total Samples	Accuracy	References
Image classification	Wheat	Varieties identification	RGB	DenseNet InceptionV3 MobileNet	31,606	95.68% 95.62% 95.49%	Laabassi et al. [114]
	Wheat	Identification of cultivars	RGB	ResNet-50 SE-ResNet SE-ResNeXt	4540	92.07%	Gao et al. [116]
	Maize	Maize seed classification	RGB	P-ResNet	8080	99.7%	Xu et al. [118]
Object detection	Maize	Classification of varieties	RGB	VGG-16	9000	98.1%	Javanmardi et al. [115]
Instance segmentation	Maize	Seed variety classification	RGB	Mask R-CNN VGG16 ResNet50 CK-CNN	16,500	64.7% 89% 92.5% 95.6%	Velesaca et al. [117]

SE—squeeze and excitation; CK—corn kernel.

4. Discussion

With the development of computer vision and deep learning, image processing has achieved great success over the last decade. One of the key techniques leading to this success is that of CNNs [121]. CNNs are often used as algorithmic tools for analyzing data. As a technique for automatic feature extraction, the CNN can be used for automatic acquisition of crop phenotype information. On the basis of that, the CNN technology, when combined with different computer vision tasks, can perform various phenotype detections of grain crops. For crop organ detection, CNN can not only count flowers, seeds, and spikes, but can also detect organ length, width, and other shape parameters, with an accuracy of over 90%. The technology can also implement effective recognition of weeds and crops by extracting leaf features. In terms of accuracy, CNNs have achieved more than 94% recognition rate for maize plants. The overall recognition rate is roughly the same as previous work when the dataset is larger, the crop growth stages are more diverse, and the background is more complex. Furthermore, CNN-based crop pest research involves a wide range of diseases, including common rusts, powdery mildew, and blast, and insect pests such as mites, wheat aphid, and corn borers. The tasks not only include basic tasks such as pest and disease classification and detection, but also refer to more complex tasks such as determination of infection levels. In addition to biotic stresses, abiotic stress phenotypes, such as nutrient deficiency and water stress assessment, showed high accuracies. It is worth mentioning that the method has potential in seed variety identification. With the limited phenotypic information of seeds, the proposed CNN models were able to classify seeds effectively with an accuracy of more than 95%. In general, CNN combined with CV has been widely used in grain crop phenotypic research.

The performance of different CNNs In the phenotype detection of grain crops is influenced by several factors. The main factor is the network architecture. Generally, deep CNN models have higher accuracy than shallow networks [20]. For example, researchers found that DenseNet-121 (with 121 layers) and ResNets (with 50 and 101 layers) had accuracies over 95% [63], while the ResNet net with 18 layers only achieved an accuracy of 88.54% for weed and crop recognitions [60]. Similarly, as an upgraded version of Faster R-CNN [85], Mask R-CNN allows simultaneous target detection, image classification and instance segmentation in a neural network due to its performance improvements in architecture. In addition to this, the performance of CNNs is also affected by the training

strategy used. Jiang et al. [77] found CNNs trained from scratch had slow convergence. In contrast, the CNNs trained by fixed feature extraction converged more rapidly but had the lowest accuracy. Furthermore, the input dataset is crucial for training CNNs, because it is the basic source of information. By providing geometrically transformed replicates of the sample images to provide a larger and more general dataset, the accuracies of CNNs were improved [55,89,90]. In addition, image quality can interfere with crop phenotyping detection results. In particular, images collected in field conditions can be affected by environmental factors such as complex backgrounds, unstable lighting, and image blur, all of which might lead to misanalysis [60,79]. Therefore, annotated datasets of large size and rich variety will always be required for CNNs. In summary, the development of different CNN model architectures with appropriate strategies and datasets can solve various phenotypic tasks. Object detection based on Fast R-CNN was more universal when crop organ was the object of counting, while Mask R-CNN showed better performance, with accuracies as high as 99%. Image classification was not widely used for the recognition of weed and crops, but it is favored by biotic and abiotic stress assessment. The performance of YOLO, GoogleNet, and Inception models were outstanding in classifying images of crops infected with pests and diseases, with accuracies of over 95% in different cases. In addition, the VGG-16 model combined with different strategies and datasets successfully completed the tasks of target detection and instance segmentation, respectively, with accuracies as high as 98% [115,117].

5. Challenges of CV and CNNs in Grain Crops and Future Trends

The annotation of datasets is a crucial factor in building robust CNN models. As CNNs need to perform different CV tasks, the relevant datasets require instance-level (bounding box) and pixel-level (mask) annotations. Both of these are very time-consuming tasks [122]. Therefore, in future work, it will be necessary to continue developing semi-supervised learning (SSL) and unsupervised learning (UL) to lower the cost and time of data labeling [123]. Advanced SSL and UL on CNN methods, such as K-means, transfer learning, and generative adversarial networkgan (GAN) [124], have permeated multiple areas of crop phenotyping study. Moreover, a number of promising approaches, such as reinforcement learning [125] and contrastive learning [126], which have succeeded in other areas to reduce the computational and energy costs, need further exploration in crop research. There is no doubt that the increasing application of advanced algorithms will effectively alleviate the problems of insufficient training data and scarcity of labeled data in grain crops.

All studies mentioned in this review used RGB images for grain crop phenotype detection. In addition to RGB cameras, more-informative sensors (e.g., multi-spectral or hyperspectral sensors) are opening new possibilities [127]. These sensors have been mounted on UAVs and autonomous robots to obtain more information, covering a larger area of crop phenotypes [123,128,129]. Especially, UAVs could help farmers to monitor their agricultural fields and apply agrochemical products to crops with ease and high precision [130]. Chemical control can be conducted more effectively.

The development of lightweight CNNs on mobile devices (e.g., cell phone and computer software) is of great practical relevance to help farmers in agricultural management. In addition, given that GPU performance on mobile devices is not inferior to GPUs on computers, processing is slower when lightweight CNNs are deployed to mobile devices. Hence, the tradeoff between accuracy, time, and memory should be considered in the model design.

It is worth mentioning that Transformer is not inferior to CNNs as another mainstream deep learning architecture in some detection studies. Compared to CNNs, Transformer has the strong advantage of a self-attention mechanism, which allows it to make exciting progress on various vision tasks, including the four tasks mentioned above, multi-modal tasks, video processing, low-level vision, and three-dimensional analysis [131,132]. Many recent studies have tried to introduce Transformer encoders into the improved model as a

convolution operation, such as to identify field crop diseases [133] and split crops of remote sensing images [134]. Numerous results show that this combined model outperforms a single CNN or Transformer approach with good generalization capabilities. This provides possibilities for transferring deep learning models to mobile phenotype detection devices. Although spectroscopic techniques based on machine learning have been extensively studied during the past few years, CV techniques based on CNNs will show greater potential in agricultural production [135–142]. Finally, it is clear that much work has been conducted on phenotyping of rice, wheat, maize, and soybean, but other species of grains and even other types of crops (e.g., fruits and vegetables) have also been explored, both to demonstrate interest in the phenotypic problem and to show the potential of CNN-based CV techniques to address it efficiently.

6. Conclusions

CNNs are widely used for phenotype detection of four grain crops. Different CNN models including VGG, YOLO, Fast R-CNN, and Mask R-CNN have been used for image classification, object detection, semantic segmentation, and instance segmentation. In this paper, we reviewed the latest CNN networks pertinent to organ counting, weed segmentation, biotic and abiotic stress assessment, and seed variety classification. The results demonstrate the importance of network architecture, development strategy, and annotated datasets in the model design for different tasks, which can directly affect the performance of CNNs. To benefit from the great potential of CNNs, high-quality sample images remain a crucial element for crop phenotyping, and robust CNNs on mobile devices are desired for practical applications. Given the recent boom in the development of CNNs combined with CV technology, it is anticipated that the method will become widespread for obtaining crop phenotype data in real time, leading to more impactful results that will contribute to precision agriculture and food security in the future.

Author Contributions: Conceptualization, W.-H.S.; methodology, W.-H.S.; investigation, Y.-H.W.; resources, W.-H.S.; writing—original draft preparation, Y.-H.W.; writing—review and editing, W.-H.S.; visualization, Y.-H.W.; supervision, W.-H.S.; project administration, W.-H.S.; funding acquisition, W.-H.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

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

References

1. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264. [CrossRef] [PubMed]
2. Steensland, A.; Thompson, T.L. *2020 Global Agricultural Productivity Report: Productivity in a Time of Pandemics. Global Agricultural Productivity Report: Productivity in a Time of Pandemics*; College of Agriculture and Life Sciences: Blacksburg, VA, USA, 2020.
3. Yu, Q.Y.; Xiang, M.T.; Wu, W.B.; Tang, H.J. Changes in global cropland area and cereal production: An inter-country comparison. *Agric. Ecosyst. Environ.* **2019**, *269*, 140–147. [CrossRef]
4. Pan, Y.H. Analysis of concepts and categories of plant phenome and phenomics. *Acta Agron. Sin.* **2015**, *41*, 175–186. [CrossRef]
5. Vithu, P.; Moses, J.A. Machine vision system for food grain quality evaluation: A review. *Trends Food Sci. Technol.* **2016**, *56*, 13–20. [CrossRef]
6. 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]
7. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. *Inf. Process. Agric.* **2021**, *8*, 27–51. [CrossRef]
8. Wang, A.C.; Zhang, W.; Wei, X.H. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.* **2019**, *158*, 226–240. [CrossRef]
9. Sun, D.; Robbins, K.; Morales, N.; Shu, Q.; Cen, H. Advances in optical phenotyping of cereal crops. *Trends Plant Sci.* **2022**, *27*, 191–208. [CrossRef] [PubMed]

10. Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [CrossRef]
11. Dhillon, A.; Verma, G.K. Convolutional neural network: A review of models, methodologies and applications to object detection. *Prog. Artif. Intell.* **2020**, *9*, 85–112. [CrossRef]
12. Mo, Y.; Wu, Y.; Yang, X.; Liu, F.; Liao, Y. Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing* **2022**, *493*, 626–646. [CrossRef]
13. 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]
14. Diba, A.; Sharma, V.; Pazandeh, A.; Pirsiavash, H.; van Gool, L. Weakly supervised cascaded convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017, Honolulu, HI, USA, 21–26 July 2017; pp. 914–922.
15. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef]
16. Liu, J.; Wang, X. Plant diseases and pests detection based on deep learning: A review. *Plant Methods* **2021**, *17*, 22. [CrossRef] [PubMed]
17. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *NIPS* **2012**, *25*, 84–90. [CrossRef]
18. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
19. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
20. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
21. Zoph, B.; Le, Q.V. Neural architecture search with reinforcement learning. *arXiv* **2016**, arXiv:1611.01578.
22. Montavon, G.; Samek, W.; Muller, K.R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [CrossRef]
23. Arrieta, A.B.; Diaz-Rodriguez, N.; del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
24. Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; LeCun, Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv* **2013**, arXiv:1312.6229.
25. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2014, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
26. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision 2015, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
27. Ren, S.Q.; He, K.M.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015.
28. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
29. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [CrossRef] [PubMed]
30. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 2015, Munich, Germany, 5–9 October 2015; pp. 234–241.
31. Hariharan, B.; Arbeláez, P.; Girshick, R.; Malik, J. Simultaneous detection and segmentation. In Proceedings of the European Conference on Computer Vision 2014, Zurich, Switzerland, 6–12 September 2014; pp. 297–312.
32. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision 2017, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
33. Jiang, Y.; Li, C. Convolutional Neural Networks for Image-Based High-Throughput Plant Phenotyping: A Review. *Plant Phenomics* **2020**, *2020*, 4152816. [CrossRef]
34. Watt, M.; Fiorani, F.; Usadel, B.; Rascher, U.; Muller, O.; Schurr, U. Phenotyping: New Windows into the Plant for Breeders. *Annu. Rev. Plant Biol.* **2020**, *71*, 689–712. [CrossRef] [PubMed]
35. Crossa, J.; Perez-Rodriguez, P.; Cuevas, J.; Montesinos-Lopez, O.; Jarquin, D.; de los Campos, G.; Burgueno, J.; Gonzalez-Camacho, J.M.; Perez-Elizalde, S.; Beyene, Y.; et al. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Sci.* **2017**, *22*, 961–975. [CrossRef] [PubMed]

36. Furbank, R.T.; Tester, M. Phenomics–technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [CrossRef] [PubMed]
37. Furbank, R.T.; Jimenez-Berni, J.A.; George-Jaeggli, B.; Potgieter, A.B.; Deery, D.M. Field crop phenomics: Enabling breeding for radiation use efficiency and biomass in cereal crops. *New Phytol.* **2019**, *223*, 1714–1727. [CrossRef]
38. Kolhar, S.; Jagtap, J. Plant trait estimation and classification studies in plant phenotyping using machine vision–A review. *Inf. Process. Agric.* **2021**. [CrossRef]
39. Kamilaris, A.; Prenafeta-Boldu, F.X. A review of the use of convolutional neural networks in agriculture. *J. Agric. Sci.* **2018**, *156*, 312–322. [CrossRef]
40. Deng, R.; Tao, M.; Huang, X.; Bangura, K.; Jiang, Q.; Jiang, Y.; Qi, L. Automated Counting Grains on the Rice Panicle Based on Deep Learning Method. *Sensors* **2021**, *21*, 281. [CrossRef]
41. Li, J.; Li, C.; Fei, S.; Ma, C.; Chen, W.; Ding, F.; Wang, Y.; Li, Y.; Shi, J.; Xiao, Z. Wheat Ear Recognition Based on RetinaNet and Transfer Learning. *Sensors* **2021**, *21*, 4845. [CrossRef]
42. Pratama, M.T.; Kim, S.; Ozawa, S.; Ohkawa, T.; Chona, Y.; Tsuji, H.; Murakami, N. Deep Learning-based Object Detection for Crop Monitoring in Soybean Fields. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN) 2020, Glasgow, UK, 19–24 July 2020; pp. 1–7.
43. Gong, B.; Ergu, D.; Cai, Y.; Ma, B. Real-Time Detection for Wheat Head Applying Deep Neural Network. *Sensors* **2020**, *21*, 191. [CrossRef] [PubMed]
44. Zou, H.; Lu, H.; Li, Y.; Liu, L.; Cao, Z. Maize tassels detection: A benchmark of the state of the art. *Plant Methods* **2020**, *16*, 108. [CrossRef] [PubMed]
45. Lu, H.; Cao, Z.; Xiao, Y.; Fang, Z.; Zhu, Y.; Xian, K. Fine-grained maize tassel trait characterization with multi-view representations. *Comput. Electron. Agric.* **2015**, *118*, 143–158. [CrossRef]
46. Lu, H.; Cao, Z.; Xiao, Y.; Li, Y.; Zhu, Y. Region-based colour modelling for joint crop and maize tassel segmentation. *Biosyst. Eng.* **2016**, *147*, 139–150. [CrossRef]
47. Sadeghi-Tehran, P.; Virlet, N.; Ampe, E.M.; Reyns, P.; Hawkesford, M.J. DeepCount: In-Field Automatic Quantification of Wheat Spikes Using Simple Linear Iterative Clustering and Deep Convolutional Neural Networks. *Front. Plant Sci.* **2019**, *10*, 1176. [CrossRef]
48. Xiong, H.; Cao, Z.; Lu, H.; Madec, S.; Liu, L.; Shen, C. TasselNetv2: In-field counting of wheat spikes with context-augmented local regression networks. *Plant Methods* **2019**, *15*, 150. [CrossRef]
49. Kienbaum, L.; Abondano, M.C.; Blas, R.; Schmid, K. DeepCob: Precise and high-throughput analysis of maize cob geometry using deep learning with an application in genebank phenomics. *Plant Methods* **2021**, *17*, 91. [CrossRef]
50. Khaki, S.; Pham, H.; Han, Y.; Kuhl, A.; Kent, W.; Wang, L.Z. 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]
51. Yang, S.; Zheng, L.; He, P.; Wu, T.; Sun, S.; Wang, M. High-throughput soybean seeds phenotyping with convolutional neural networks and transfer learning. *Plant Methods* **2021**, *17*, 50. [CrossRef]
52. Machefer, M.; Lemarchand, F.; Bonnefond, V.; Hitchins, A.; Sidiropoulos, P. Mask R-CNN refitting strategy for plant counting and sizing in UAV imagery. *Remote Sens.* **2020**, *12*, 3015. [CrossRef]
53. Li, S.; Yan, Z.; Guo, Y.; Su, X.; Cao, Y.; Jiang, B.; Yang, F.; Zhang, Z.; Xin, D.; Chen, Q.; et al. SPM-IS: An auto-algorithm to acquire a mature soybean phenotype based on instance segmentation. *Crop J.* **2021**, *10*, 1412–1423. [CrossRef]
54. Tan, C.; Li, C.; He, D.; Song, H. Anchor-free deep convolutional neural network for plant and plant organ detection and counting. In Proceedings of the 2021 ASABE Annual International Virtual Meeting, Online, 12–16 July 2021; p. 1. [CrossRef]
55. Li, Y.; Jia, J.D.; Zhang, L.; Khattak, A.M.; Sun, S.; Gao, W.L.; Wang, M.J. Soybean Seed Counting Based on Pod Image Using Two-Column Convolution Neural Network. *IEEE Access* **2019**, *7*, 64177–64185. [CrossRef]
56. Ying, W.; Yue, L.; Tingting, W.; Shi, S.; Minjuan, W. Fast Counting Method of Soybean Seeds Based on Density Estimation and VGG-Two. *Smart Agric.* **2021**, *3*, 111.
57. Korav, S.; Dhaka, A.K.; Singh, R.; Premaradhya, N.; Reddy, G.C. A study on crop weed competition in field crops. *J. Pharmacogn. Phytochem.* **2018**, *7*, 3235–3240.
58. Agrawal, K.N.; Singh, K.; Bora, G.C.; Lin, D. Weed recognition using image-processing technique based on leaf parameters. *J. Agric. Sci. Technol. B* **2012**, *2*, 899.
59. Lin, S.M.; Jiang, Y.; Chen, X.S.; Biswas, A.; Li, S.; Yuan, Z.H.; Wang, H.L.; Qi, L. Automatic Detection of Plant Rows for a Transplanter in Paddy Field Using Faster R-CNN. *IEEE Access* **2020**, *8*, 147231–147240. [CrossRef]
60. Wang, S.S.; Zhang, W.Y.; Wang, X.S.; Yu, S.S. Recognition of rice seedling rows based on row vector grid classification. *Comput. Electron. Agric.* **2021**, *190*, 106454. [CrossRef]
61. Quan, L.; Feng, H.; Lv, Y.; Wang, Q.; Zhang, C.; Liu, J.; Yuan, Z. Maize seedling detection under different growth stages and complex field environments based on an improved Faster R–CNN. *Biosyst. Eng.* **2019**, *184*, 1–23. [CrossRef]
62. 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]
63. Kim, Y.H.; Park, K.R. MTS-CNN: Multi-task semantic segmentation-convolutional neural network for detecting crops and weeds. *Comput. Electron. Agric.* **2022**, *199*, 107146. [CrossRef]

64. Weirong, Z.; Haojun, W.; Chaofan, Q.; Guangyan, W. Maize Seedling and Core Detection Method Based on Mask R-CNN. *Xinjiang Agric. Sci.* **2021**, *58*, 1918–1928.

65. Zhang, R.; Wang, C.; Hu, X.; Liu, Y.; Chen, S. Weed location and recognition based on UAV imaging and deep learning. *Int. J. Precis. Agric. Aviat.* **2020**, *3*, 23–29. [CrossRef]

66. Haq, M.A. CNN Based Automated Weed Detection System Using UAV Imagery. *Comput. Syst. Sci. Eng.* **2022**, *42*, 837–849.

67. Babu, V.S.; Ram, N.V. Deep Residual CNN with Contrast Limited Adaptive Histogram Equalization for Weed Detection in Soybean Crops. *Traitement Signal* **2022**, *39*, 717–722. [CrossRef]

68. 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]

69. 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] [PubMed]

70. Zhang, J.-L.; Su, W.-H.; Zhang, H.-Y.; Peng, Y. SE-YOLOv5x: An Optimized Model Based on Transfer Learning and Visual Attention Mechanism for Identifying and Localizing Weeds and Vegetables. *Agronomy* **2022**, *12*, 2061. [CrossRef]

71. Zhang, Z.Z.; Kayacan, E.; Thompson, B.; Chowdhary, G. High precision control and deep learning-based corn stand counting algorithms for agricultural robot. *Auton. Robot.* **2020**, *44*, 1289–1302. [CrossRef]

72. Fina, F.; Birch, P.; Young, R.; Obu, J.; Faithpraise, B.; Chatwin, C. Research. Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters. *Int. J. Adv. Biotechnol. Res.* **2013**, *4*, 189–199.

73. Sharma, R.; Kukreja, V.; Kadyan, V. Hispa Rice Disease Classification using Convolutional Neural Network. In Proceedings of the 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Tamil Nadu, India, 13–14 May 2021; pp. 377–381.

74. Krishnamoorthy, N.; Prasad, L.V.N.; Kumar, C.S.P.; Subedi, B.; Abraha, H.B.; Sathishkumar, V.E. Rice leaf diseases prediction using deep neural networks with transfer learning. *Environ. Res.* **2021**, *198*, 111275. [CrossRef]

75. Singh, A.; Arora, M. CNN Based Detection of Healthy and Unhealthy Wheat Crop. In Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 10–12 September 2020; pp. 121–125.

76. Kumar, D.; Kukreja, V. N-CNN Based Transfer Learning Method for Classification of Powdery Mildew Wheat Disease. In Proceedings of the 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 5–7 March 2021; pp. 707–710.

77. Jiang, J.L.; Liu, H.Y.; Zhao, C.; He, C.; Ma, J.F.; Cheng, T.; Zhu, Y.; Cao, W.X.; Yao, X. Evaluation of Diverse Convolutional Neural Networks and Training Strategies for Wheat Leaf Disease Identification with Field-Acquired Photographs. *Remote Sens.* **2022**, *14*, 3446. [CrossRef]

78. Bari, B.S.; Islam, M.N.; Rashid, M.; Hasan, M.J.; Razman, M.A.M.; Musa, R.M.; Ab Nasir, A.F.; Majeed, A.P.P.A. A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework. *PeerJ Comput. Sci.* **2021**, *7*, e432. [CrossRef]

79. Zhou, G.X.; Zhang, W.Z.; Chen, A.B.; He, M.F.; Ma, X.S. Rapid Detection of Rice Disease Based on FCM-KM and Faster R-CNN Fusion. *IEEE Access* **2019**, *7*, 143190–143206. [CrossRef]

80. Zhang, K.K.; Wu, Q.F.; Chen, Y.P. Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN. *Comput. Electron. Agric.* **2021**, *183*, 106064. [CrossRef]

81. Shrivastava, S.; Singh, S.K.; Hooda, D.S. Soybean plant foliar disease detection using image retrieval approaches. *Multimed. Tools Appl.* **2017**, *76*, 26647–26674. [CrossRef]

82. Pires, R.D.L.; Gonçalves, D.N.; Oruê, J.P.M.; Kanashiro, W.E.S.; Rodrigues, J.F., Jr.; Machado, B.B.; Gonçalves, W.N. Local descriptors for soybean disease recognition. *Comput. Electron. Agric.* **2016**, *125*, 48–55. [CrossRef]

83. Ennadifi, E.; Laraba, S.; Vincke, D.; Mercatoris, B.; Gosselin, B. Wheat Diseases Classification and Localization Using Convolutional Neural Networks and GradCAM Visualization. In Proceedings of the 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 9–11 June 2020; pp. 1–5.

84. Lin, Z.Q.; Mu, S.M.; Huang, F.; Mateen, K.A.; Wang, M.J.; Gao, W.L.; Jia, J.D. A Unified Matrix-Based Convolutional Neural Network for Fine-Grained Image Classification of Wheat Leaf Diseases. *IEEE Access* **2019**, *7*, 11570–11590. [CrossRef]

85. Su, W.-H.; Zhang, J.; Yang, C.; Page, R.; Szinyei, T.; Hirsch, C.D.; Steffenson, B. Automatic evaluation of wheat resistance to fusarium head blight using dual mask-RCNN deep learning frameworks in computer vision. *Remote Sens.* **2020**, *13*, 26. [CrossRef]

86. Gao, Y.; Wang, H.; Li, M.; Su, W.-H. Automatic Tandem Dual BlendMask Networks for Severity Assessment of Wheat Fusarium Head Blight. *Agriculture* **2022**, *12*, 1493. [CrossRef]

87. Krishnamoorthi, M.; Sankavi, R.S.; Aishwarya, V.; Chithra, B. Maize Leaf Diseases Identification using Data Augmentation and Convolutional Neural Network. In Proceedings of the 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 7–9 October 2021; pp. 1672–1677.

88. Zhang, Y.; Wa, S.Y.; Liu, Y.T.; Zhou, X.Y.; Sun, P.S.; Ma, Q. High-Accuracy Detection of Maize Leaf Diseases CNN Based on Multi-Pathway Activation Function Module. *Remote Sens.* **2021**, *13*, 4218. [CrossRef]

89. Hasan, M.J.; Alom, M.S.; Dina, U.F.; Moon, M.H. Maize Diseases Image Identification and Classification by Combining CNN with Bi-Directional Long Short-Term Memory Model. In Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, 5–7 June 2020; pp. 1804–1807.

90.   Arora, J.; Agrawal, U. Systems. Classification of Maize leaf diseases from healthy leaves using Deep Forest. *J. Artif. Intell. Syst.* **2020**, *2*, 14–26. [CrossRef]

91.   Karlekar, A.; Seal, A. SoyNet: Soybean leaf diseases classification. *Comput. Electron. Agric.* **2020**, *172*, 105342. [CrossRef]

92.   Bao, W.X.; Yang, X.H.; Liang, D.; Hu, G.S.; Yang, X.J. Lightweight convolutional neural network model for field wheat ear disease identification. *Comput. Electron. Agric.* **2021**, *189*, 106367. [CrossRef]

93.   Pan, Q.; Gao, M.; Wu, P.; Yan, J.; Li, S. A Deep-Learning-Based Approach for Wheat Yellow Rust Disease Recognition from Unmanned Aerial Vehicle Images. *Sensors* **2021**, *21*, 6540. [CrossRef] [PubMed]

94.   Baliyan, A.; Kukreja, V.; Salonki, V.; Kaswan, K.S. Detection of Corn Gray Leaf Spot Severity Levels using Deep Learning Approach. In Proceedings of the 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 3–4 September 2021; pp. 1–5.

95.   Samanta, R.; Ghosh, I. Tea insect pests classification based on artificial neural networks. *Int. J. Comput. Eng. Sci.* **2012**, *2*, 1–13.

96.   Geiger, F.; Bengtsson, J.; Berendse, F.; Weisser, W.W.; Emmerson, M.; Morales, M.B.; Ceryngier, P.; Liira, J.; Tscharntke, T.; Winqvist, C.; et al. Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic Appl. Ecol.* **2010**, *11*, 97–105. [CrossRef]

97.   Clement, A.; Verfaille, T.; Lormel, C.; Jaloux, B. A new colour vision system to quantify automatically foliar discolouration caused by insect pests feeding on leaf cells. *Biosyst. Eng.* **2015**, *133*, 128–140. [CrossRef]

98.   Liu, T.; Chen, W.; Wu, W.; Sun, C.; Guo, W.; Zhu, X. Detection of aphids in wheat fields using a computer vision technique. *Biosyst. Eng.* **2016**, *141*, 82–93. [CrossRef]

99.   Ishengoma, F.S.; Rai, I.A.; Said, R.N. Identification of maize leaves infected by fall armyworms using UAV-based imagery and convolutional neural networks. *Comput. Electron. Agric.* **2021**, *184*, 106124. [CrossRef]

100. Tetila, E.C.; Machado, B.B.; Menezes, G.V.; de Belete, N.A.S.; Astolfi, G.; Pistori, H. A deep-learning approach for automatic counting of soybean insect pests. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1837–1841. [CrossRef]

101. Abade, A.; Porto, L.F.; Ferreira, P.A.; de Vidal, F.B. NemaNet: A convolutional neural network model for identification of soybean nematodes. *Biosyst. Eng.* **2022**, *213*, 39–62. [CrossRef]

102. Li, R.; Wang, R.J.; Zhang, J.; Xie, C.J.; Liu, L.; Wang, F.Y.; Chen, H.B.; Chen, T.J.; Hu, H.Y.; Jia, X.F.; et al. An Effective Data Augmentation Strategy for CNN-Based Pest Localization and Recognition in the Field. *IEEE Access* **2019**, *7*, 160274–160283. [CrossRef]

103. Sheema, D.; Ramesh, K.; Renjith, P.N.; Lakshna, A. Comparative Study of Major Algorithms for Pest Detection in Maize Crop. In Proceedings of the 2021 International Conference on Intelligent Technologies (CONIT), Hubli, India, 25–27 June 2021; pp. 1–7.

104. Verma, S.; Tripathi, S.; Singh, A.; Ojha, M.; Saxena, R.R. Insect Detection and Identification using YOLO Algorithms on Soybean Crop. In Proceedings of the TENCON 2021–2021 IEEE Region 10 Conference (TENCON), Auckland, New Zealand, 7–10 December 2021; pp. 272–277.

105. Chen, P.; Li, W.L.; Yao, S.J.; Ma, C.; Zhang, J.; Wang, B.; Zheng, C.H.; Xie, C.J.; Liang, D. Recognition and counting of wheat mites in wheat fields by a three-step deep learning method. *Neurocomputing* **2021**, *437*, 21–30. [CrossRef]

106. Anwar, A.; Kim, J.K. Transgenic Breeding Approaches for Improving Abiotic Stress Tolerance: Recent Progress and Future Perspectives. *Int. J. Mol. Sci.* **2020**, *21*, 2695. [CrossRef] [PubMed]

107. 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] [PubMed]

108. Sethy, P.K.; Barpanda, N.K.; Rath, A.K.; Behera, S.K. Nitrogen Deficiency Prediction of Rice Crop Based on Convolutional Neural Network. *J. Ambient. Intell. Humaniz. Comput.* **2020**, *11*, 5703–5711. [CrossRef]

109. Wang, C.; Ye, Y.; Tian, Y.; Yu, Z. Classification of nutrient deficiency in rice based on CNN model with Reinforcement Learning augmentation. In Proceedings of the 2021 International Symposium on Artificial Intelligence and Its Application on Media (ISAIAM), Xi'an, China, 21–23 May 2021; pp. 107–111.

110. Rizal, S.; Pratiwi, N.K.C.; Ibrahim, N.; Syalomta, N.; Nasution, M.I.K.; Mz, I.M.U.; Oktavia, D.A.P. Classification Of Nutrition Deficiency In Rice Plant Using CNN. In Proceedings of the 2022 1st International Conference on Information System & Information Technology (ICISIT), Yogyakarta, Indonesia, 26–27 July 2022; pp. 382–385.

111. Zhuang, S.; Wang, P.; Jiang, B.R.; Li, M.S. Learned features of leaf phenotype to monitor maize water status in the fields. *Comput. Electron. Agric.* **2020**, *172*, 105347. [CrossRef]

112. Das, S.; Christopher, J.; Apan, A.; Choudhury, M.R.; Chapman, S.; Menzies, N.W.; Dang, Y.P. Evaluation of water status of wheat genotypes to aid prediction of yield on sodic soils using UAV-thermal imaging and machine learning. *Agric. For. Meteorol.* **2021**, *307*, 108477. [CrossRef]

113. Shouche, S.P.; Rastogi, R.; Bhagwat, S.G.; Sainis, J.K. Shape analysis of grains of Indian wheat varieties. *Comput. Electron. Agric.* **2001**, *33*, 55–76. [CrossRef]

114. Laabassi, K.; Belarbi, M.A.; Mahmoudi, S.; Mahmoudi, S.A.; Ferhat, K. Wheat varieties identification based on a deep learning approach. *J. Saudi Soc. Agric. Sci.* **2021**, *20*, 281–289. [CrossRef]

115. Javanmardi, S.; Ashtiani, S.H.M.; Verbeek, F.J.; Martynenko, A. Computer-vision classification of corn seed varieties using deep convolutional neural network. *J. Stored Prod. Res.* **2021**, *92*, 101800. [CrossRef]

116. Gao, J.; Liu, C.; Han, J.; Lu, Q.; Wang, H.; Zhang, J.; Bai, X.; Luo, J. Identification Method of Wheat Cultivars by Using a Convolutional Neural Network Combined with Images of Multiple Growth Periods of Wheat. *Symmetry* **2021**, *13*, 2012. [CrossRef]

117. Velesaca, H.O.; Mira, R.; Suárez, P.L.; Larrea, C.X.; Sappa, A.D. Deep learning based corn kernel classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2020, Seattle, WA, USA, 14–19 June 2020; pp. 66–67.

118. Xu, P.; Tan, Q.; Zhang, Y.; Zha, X.; Yang, S.; Yang, R. Research on Maize Seed Classification and Recognition Based on Machine Vision and Deep Learning. *Agriculture* **2022**, *12*, 232. [CrossRef]

119. ElMasry, G.; ElGamal, R.; Mandour, N.; Gou, P.; Al-Rejaie, S.; Belin, E.; Rousseau, D. Emerging thermal imaging techniques for seed quality evaluation: Principles and applications. *Food Res. Int.* **2020**, *131*, 109025. [CrossRef]

120. Zhu, S.; Zhang, J.; Chao, M.; Xu, X.; Song, P.; Zhang, J.; Huang, Z. A Rapid and Highly Efficient Method for the Identification of Soybean Seed Varieties: Hyperspectral Images Combined with Transfer Learning. *Molecules* **2019**, *25*, 152. [CrossRef]

121. Khosrokhani, M.; Nasr, A.H. Applications of the Remote Sensing Technology to Detect and Monitor the Rust Disease in the Wheat–A Literature Review. *Geocarto Int.* 2022, 1–27, *accepted*. [CrossRef]

122. Murthy, V.N.; Maji, S.; Manmatha, R. Automatic image annotation using deep learning representations. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, Shanghai, China, 23–26 June 2015; pp. 603–606.

123. Blok, P.M.; Kootstra, G.; Elghor, H.E.; Diallo, B.; van Evert, F.K.; van Henten, E.J. Active learning with MaskAL reduces annotation effort for training Mask R-CNN. *Comput. Electron. Agric.* **2021**, *197*, 106917. [CrossRef]

124. Li, J.; Jia, J.; Xu, D. Unsupervised representation learning of image-based plant disease with deep convolutional generative adversarial networks. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018; pp. 9159–9163.

125. Eckardt, J.-N.; Wendt, K.; Bornhäuser, M.; Middeke, J.M. Reinforcement learning for precision oncology. *Cancers* **2021**, *13*, 4624. [CrossRef]

126. Wang, X.; Qi, G.-J. Contrastive learning with stronger augmentations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, 1–12. [CrossRef] [PubMed]

127. Khaled, A.Y.; Aziz, S.A.; Bejo, S.K.; Nawi, N.M.; Seman, I.A.; Onwude, D.I. Early detection of diseases in plant tissue using spectroscopy–applications and limitations. *Appl. Spectrosc. Rev.* **2018**, *53*, 36–64. [CrossRef]

128. Zhang, S.M.; Li, X.H.; Ba, Y.X.; Lyu, X.G.; Zhang, M.Q.; Li, M.Z. Banana Fusarium Wilt Disease Detection by Supervised and Unsupervised Methods from UAV-Based Multispectral Imagery. *Remote Sens.* **2022**, *14*, 1231. [CrossRef]

129. Allmendinger, A.; Spaeth, M.; Saile, M.; Peteinatos, G.G.; Gerhards, R. Precision Chemical Weed Management Strategies: A Review and a Design of a New CNN-Based Modular Spot Sprayer. *Agronomy* **2022**, *12*, 1620. [CrossRef]

130. Bouguettaya, A.; Zarzour, H.; Kechida, A.; Taberkit, A.M. Recent Advances on UAV and Deep Learning for Early Crop Diseases Identification: A Short Review. In Proceedings of the 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 14–15 July 2021; pp. 334–339.

131. Ang, K.L.-M.; Seng, J.K.P. Big data and machine learning with hyperspectral information in agriculture. *IEEE Access* **2021**, *9*, 36699–36718. [CrossRef]

132. Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. *ACM Comput. Surv.* **2022**, *54*, 1–41. [CrossRef]

133. Zhu, W.; Sun, J.; Wang, S.; Shen, J.; Yang, K.; Zhou, X. Identifying Field Crop Diseases Using Transformer-Embedded Convolutional Neural Network. *Agriculture* **2022**, *12*, 1083. [CrossRef]

134. Wang, H.; Chen, X.; Zhang, T.; Xu, Z.; Li, J. CCTNet: Coupled CNN and Transformer Network for Crop Segmentation of Remote Sensing Images. *Remote Sens.* **2022**, *14*, 1956. [CrossRef]

135. Su, W.-H.; He, H.-J.; Sun, D.-W. Non-destructive and rapid evaluation of staple foods quality by using spectroscopic techniques: A review. *Crit. Rev. Food Sci. Nutr.* **2017**, *57*, 1039–1051. [CrossRef] [PubMed]

136. Su, W.-H.; Bakalis, S.; Sun, D.-W. Fingerprinting study of tuber ultimate compressive strength at different microwave drying times using mid-infrared imaging spectroscopy. *Dry. Technol.* **2019**, *37*, 1113–1130. [CrossRef]

137. Liu, B.-Y.; Fan, K.-J.; Su, W.-H.; Peng, Y. Two-Stage Convolutional Neural Networks for Diagnosing the Severity of Alternaria Leaf Blotch Disease of the Apple Tree. *Remote Sens.* **2022**, *14*, 2519. [CrossRef]

138. Su, W.-H.; Xue, H. Imaging Spectroscopy and Machine Learning for Intelligent Determination of Potato and Sweet Potato Quality. *Foods* **2021**, *10*, 2146. [CrossRef]

139. Fan, K.J.; Su, W.H. Applications of Fluorescence Spectroscopy, RGB-and Multispectral Imaging for Quality Determinations of White Meat: A Review. *Biosensors* **2022**, *12*, 76. [CrossRef]

140. Su, W.-H.; Sheng, J.; Huang, Q.-Y. Development of a Three-Dimensional Plant Localization Technique for Automatic Differentiation of Soybean from Intra-Row Weeds. *Agriculture* **2022**, *12*, 195. [CrossRef]

141. Su, W.-H.; Sun, D.-W. Advanced analysis of roots and tubers by hyperspectral techniques. In *Advances in Food and Nutrition Research*; Fidel, T., Ed.; Elsevier: Amsterdam, The Netherlands, 2019; Volume 87, pp. 255–303.

142. Su, W.-H. Advanced Machine Learning in Point Spectroscopy, RGB-and hyperspectral-imaging for automatic discriminations of crops and weeds: A review. *Smart Cities* **2020**, *3*, 39. [CrossRef]