applied sciences

Article

A Lightweight Neural Network-Based Method for Identifying Early-Blight and Late-Blight Leaves of Potato

Feilong Kang [1], Jia Li [1,2,*], Chunguang Wang [1] and Fuxiang Wang [1]

[1] Inner Mongolia Agricultural University, Hohhot 010018, China
[2] Inner Mongolia Autonomous Region Key Laboratory of Big Data Research and Application of Agriculture and Animal Husbandry, Hohhot 010018, China
* Correspondence: lijia@imau.edu.cn

Abstract: Crop pests and diseases are one of the most critical disasters that limit agricultural production. In this paper, we trained a lightweight convolutional neural network model and built a Django framework-based potato disease leaf recognition system, which can recognize three types of potato leaf images including early blight, late blight, and healthy. A lightweight, neural network-based model for the identification of early potato leaf diseases significantly reduces the number of model parameters, whereas the accuracy of Top-1 identification is over 93%. We imported the trained model into the Django framework to build a website for a potato early leaf disease identification system, thus providing technical support for the implementation of a mobile-based potato leaf disease identification and early warning system.

Keywords: convolutional neural networks; machine learning; potato disease leaf; Django framework

Citation: Kang, F.; Li, J.; Wang, C.; Wang, F. A Lightweight Neural Network-Based Method for Identifying Early-Blight and Late-Blight Leaves of Potato. *Appl. Sci.* **2023**, *13*, 1487. https://doi.org/10.3390/app13031487

Academic Editors: Junchi Yan and Minghao Guo

Received: 25 November 2022
Revised: 18 January 2023
Accepted: 20 January 2023
Published: 23 January 2023

1. Introduction

Crop pests and diseases are one of the critical disasters that limit agricultural production, directly affecting the healthy development of agricultural economies and income of farmers. In recent years, crop pests and diseases have become more frequent, and major crop pests and diseases occur from time to time. Non-destructive detection and early identification of crop pests and diseases is the key to the development of both precision agriculture and ecological agriculture. Potato is one of the most important food crops. Potatoes are highly susceptible to fungal and bacterial infections. The diversity and uncertainty of diseases make it difficult to predict, diagnose, and control its occurrence. Outbreaks often have devastating consequences and cause serious losses, becoming one of the major disasters that limit potato production.

When potatoes are affected by a variety of typical diseases such as fungi and bacteria, most of them start on the foliage and spread to the entire plant which causes severe disease and damage to the potato block. Some diseases can first damage the tubers and then spread to the leaves, causing the entire plant to wilt. Regardless of where the disease first appears, it is important to take action as soon as a lesion is observed on the leaves in order to control the spread of fungal and bacterial diseases and protect the potato crop. The implementation of precision and intelligent agriculture practices is beneficial in promoting the development of sustainable agriculture, guaranteeing food safety and protecting human health, and holds significant social and economic value.

Traditional methods of diagnosing pests and diseases in crops involve taking samples from the field and analyzing them chemically, which can be time-consuming, labor-intensive, slow to provide a diagnosis, and have a limited scope of detection. Many scholars mainly use visible imaging techniques, infrared imaging techniques, or hyperspectral imaging techniques [1] to acquire image photographs. Classical image processing techniques (including denoising, erosion and expansion, image segmentation, image enhancement, feature extraction) and machine-learning techniques (including k-mean clustering algorithms, support vector machines, artificial neural networks) are used to detect changes in the external characteristics, internal chemical content, and physiological structure of the subject for crop disease diagnosis. Many scholars have used classical image processing techniques and machine-learning methods to diagnose and identify different crop pest targets based on acquired infrared images or hyperspectral images [2]. Chen et al. used convolution network architecture to recognize potato diseases [3]. Raza et al. used a combination of infrared image information and depth information to achieve detection of tomato powdery mildew using SVM [4]. Zhao et al. [5] and Pandey et al. [6] used machine learning methods to analyze hyperspectral images of plant leaves such as cucumber.

Using classical machine-learning classification methods to identify some specific image features of crops can face some difficulties. First, classical image-processing methods are difficult to extract similar image feature values under different lighting conditions, thus failing to accurately identify disease image targets. Second, the image target size, feature color, and regional distribution of crop diseases at different onset stages are not consistent, and it is difficult to represent them using multiple features within a certain range. Some characteristics, such as those resembling the color of the leaf background, make it difficult for traditional methods to effectively and accurately identify pests in images. This makes it difficult for classical image processing techniques based on morphological segmentation to extract targets effectively, which can interfere with the target recognition results.

With the rise of deep-learning techniques, more and more researchers are introducing them into the field of crop pest and disease image detection [7]. This new nondestructive detection technique can automatically extract features of crop pest images in the visible range and a achieve rapid nondestructive identification [8] without the use of hyperspectral imaging techniques, with higher accuracy, faster detection speed and better stability [9]. The use of artificial intelligence technologies such as deep learning to identify potato diseases early and take professional control measures can not only reduce the scale of the disease and ensure potato production, but also effectively reduce the amount of pesticides used. Reducing the harm caused by pesticides is an inevitable trend in the development of intelligent and green potato disease control.

Deep learning has a more complex network structure, stronger learning ability, and higher recognition accuracy than classical neural networks, and has obvious advantages in solving image classification and visualization problems [10]. In recent years, scholars have generally paid high attention to the research of deep learning; the mature deep-learning networks include restricted Boltzmann machine (RBM), deep confidence network (DBN), deep Boltzmann machine (DBM), convolutional neural network (CNN), recurrent neural network (RNN), generative adversarial network (GAN) [11], and capsule network (CapsNet). The technology of convolutional neural networks is developing rapidly, and a large number of high-performing neural network models with have emerged, such as LeNet-5, AlexNet, ZF-Net, VGGNet, GoogLeNet, ResNet, R-CNN, Faster R-CNN, YOLO, SSD [12], and DenseNet [13], etc. These models have extraordinary performance in the fields of facial recognition, vehicle detection, speech recognition, and poetry composition [14].

More and more scholars are applying deep learning to the diagnosis and identification of crop pests and diseases [15], and the use of transfer learning allows the creation of network models suitable for dedicated datasets based on existing mature convolutional neural networks [16]. Yong Zhong et al. proposed three regression, multi-label classification, and focal loss function approaches to identify apple leaf diseases based on DenseNet-121 deep convolutional network [17]. Wu et al. proposed compression and splitting techniques for image classification as an effective method to improve feature extraction ability and recognition performance for fine-grained image classification [18]. Amara et al. used LeNet convolutional neural network for banana leaf disease recognition [19]. Mohanty et al. trained a model using the GoogLeNet convolutional neural network to analyze 54,306 images of plant leaves for 26 types of diseases of 14 crops by using migration learning [20]. Amanda et al. applied migration learning to build an Inception v3 convolutional neural network model to identify cassava pests and diseases [21]. Sladojevic et al. developed a CaffeNet convolutional neural network-based model on the Caffe framework to identify leaf data of 13 crop pests and diseases [22]. Brahimi et al. compared AlexNet and GoogLeNet for the recognition of nine pests on 14,828 tomato leaves [23]. The GoogLeNet algorithm increases the width of the network and reduces the depth of the network through the Inception structure, which can reduce the amount of parameter computation and ensure the performance of the network recognition. DenseNet, on the other hand, connects all layers. The network architecture is designed to be narrower with fewer parameters, which can effectively alleviate the gradient disappearance problem due to back-propagation training of the deep network.

Initial exploration of crop pest diagnosis using classical models has achieved more satisfactory recognition results under laboratory conditions, but the existing deep-learning models for crop pest images have deep layers, complex network structures, and require a very large number of model parameters for training, which remains far from being practical. The widely used deep learning neural network technology uses ResNet as the base network, which has the disadvantage of the number of model parameters and computation being too large. Many of the deep learning models used have more than 10 million parameters, and the heavy computational load makes it difficult to deploy in mobile applications. Therefore, in order to improve the practicality, it becomes extremely important to optimize the model of convolutional neural networks.

2. Materials and Methods

We investigate the feature detection method of leaf image and disease species identification based on multi-scale pyramid fusion technology. The fusion of multi-scale features of potato leaf images has been achieved by comparing and analyzing the improved backbone network, optimizing the selection of different scale convolutional feature detectors, and using up and down sampling techniques. By analyzing deep neural network lightweighting methods such as the inverse residual technique, deep separable convolution technique and model compression technique to reduce the number of recognition model parameters, including deep convolution, point-by-point convolution and bottleneck, we provide technical support to realize a mobile system for early recognition of potato disease leaf images.

2.1. Data Set and Data Augmentation Methods

In this paper, a total of 5450 images of three types of potato leaves including early blight, late blight, and healthy were collected and labeled. The training set, validation set, and test set were set in the ratio of 7:2:1, with 3815 images in the training set, 1090 images in the validation set, and 545 images in the test set. Figure 1 shows an example diagram of a potato leaf disease. The data augmentation technique adds images of many scenes and enhances the multidimensional changes of the input images, including random scaling, cropping, flipping, rotating, and adjusting the brightness and contrast, etc. A data enhancement method based on the Mosaic algorithm is used to blend and superimpose different photos into new images as samples. The sample data photos are synthesized to generate many data samples with complex backgrounds, which enables the trained detection model to obtain higher robustness.

Figure 1. Example of potato leaf diseases. Three types of potato leaves including early blight, late blight, and healthy.

2.2. Deep Learning Models

2.2.1. ResNet Model

The residual network (ResNet model) solves the problem of difficult in the training of deep networks by fitting residual representations with shortcut connections, which effectively alleviates the problem of gradient disappearance and network degradation. This method allows the network to reach tens or even hundreds of layers, extending the number of layers of the network to a very large scale. For example, the 152-layer residual network is eight times deeper than the VGG network, but its complexity is lower. It achieved a top-five error rate of 3.57% on the ImageNet test set. The residual network has been widely used in video classification, face recognition, garbage classification, etc. The commonly residual network models include ResNet18, ResNet50, ResNet101, etc.

In this paper, the input network size of potato leaf disease image recognition model is set to 256×256, and the 64×64 size feature map is output after a convolution operation with a 7×7 size and s = 2 convolution kernel, followed by a 3×3 convolution kernel, and s = 2 pooling layer down sampling operation. The feature maps are transformed into 2048-dimensional vectors using a bottleneck residual structure and global average pooling, and the probability values of each category are output by changing the output layer and SoftMax function to complete the classification recognition of the input leaf disease images.

2.2.2. Xception Model

The Xception architecture is a linear stack of deeply separable convolutional layers with residual connections, and the modules are connected by linear residuals. In this paper, the model decomposes the scalar convolution into spatial convolution and point-by-point convolution and performs 3×3 convolution operations for each channel after 1×1 convolution, respectively, which is finally expressed by a fully connected layer. By separating the tasks of learning spatial correlation and learning inter-channel correlation, the theoretical computation of the model is significantly reduced and only a small amount of accuracy is lost.

2.2.3. MobileNet Model

MobileNetV1 [24] was proposed by Google to focus on lightweight neural networks on mobile devices, and its main contribution was to replace the standard convolutional layers in VGG with deeply separable convolutions. Since then, lightweight neural networks started to make great progress, and soon MobileNetV2 [25], MobileNetV3 [26], and other types of lightweight neural networks were born. The MobileNet series models uses deep separable convolution techniques to significantly reduce the number of model parameters while still achieving similar recognition accuracy. The core idea of MobileNetV3 is to complete the model formation with Neural Architecture Search (NAS). The model introduces deep separable convolution, a linear BottleNeck's inverse residual structure, a lightweight attention model based on squeeze and excitation structure, and a new activation function h-swish(x).

MobileNet is a lightweight network with features such as reduced model computational complexity and reduced model size. Its basic unit is the depth-separable convolution, which can be split into two smaller operations: deep convolution and point-by-point convolution. The difference between depth convolution and standard convolutional network is that the convolution kernel is split into a single channel form, and the convolution operation is performed on each channel without changing the depth of the input feature image, so that the output feature map with the same number of channels as the input feature map is obtained. Point-by-point convolution is 1×1 convolution, and its main function is to up-dimension or down-dimension the feature map, to obtain more high-dimensional spatial information. If the size of the standard convolution is set to $D_k * D_k * M$, and M is the number of convolution channels. Then the size of the convolution kernel after the convolution of N kernels is $D_w * D_h$. The number of parameters of the convolution operation is $D_k * D_k * M * N$, and the computation is $D_k * D_k * M * N * D_w * D_h$. If the depth-separable convolution is used, the number of parameters becomes $D_k * D_k * M + M * N$, and the computation is $D_k * D_k * M * D_w * D_h + M * N * D_w * D_h$. After comparison, it can be concluded that the number of parameters and the amount of computation are reduced to $\frac{1}{N} * \frac{1}{D_k^2}$. If we take the commonly used 3*3-convolution kernel as an example, it can be reduced to about one-ninth of the original size. Deep separable convolution achieves the construction of lightweight networks with fewer parameters and fewer operations. However, deep convolution does not have the ability to change channels and can only work in low dimensions, so MobileNetV2 uses the inverse residual technique, which means that point-by-point convolution is used to perform the up-dimensional operation first, and then extract features by convolution in a higher dimensional space, and then output the result by point-by-point convolution to downscale afterwards. This is the opposite of ResNet's method of first descending convolution and then ascending. In order to solve the problem of ReLU activation function generating a large number of 0 parameters, the linear bottleneck structure of linear activation function is used.

Keras is a very easy-to-use neural network library based on the tensorflow platform. Keras provides most of the modules needed to build deep learning models. Developers can easily implement research ideas using transfer learning methods. Transfer learning is typically used for tasks with small-scale datasets. Feature extraction methods learned on existing large-scale datasets are used to solve new target recognition problems. Fine-tuning can potentially achieve meaningful improvements, by incrementally adapting the pretrained features to the new data. In this paper, the model first performs a standard 3×3 convolution operation on the input potato leaf disease images, then performs a depth-separable convolution operation, changing the number of feature channels by adjusting the multiplication factor, separating the depth convolution from point-by-point convolution, and part of the depth convolution will be down-sampled by stride = 2. Then average pooling is used to turn the features into 1×1. Finally, the results are inputted to the SoftMax layer by adding a fully connected layer according to categories, thus constructing a potato leaf disease recognition model.

3. Results

The number of parameters of a network model determines the complexity of its computation. The huge amount of computation leads to the existing mature deep learning models cannot be directly applied to mobile. Therefore, researching lightweight neural network structures is the first task in the goal of reducing the number of model parameters. If some modules of the deep neural network model are directly reduced, it will definitely affect the accuracy of potato leaf image feature extraction and lead to the decrease of potato disease recognition model accuracy. Based on the above analysis of the residual network structure and the depth-separable convolutional network model, the number of model parameters is reduced by using the depth-separable convolutional technique for the characteristics of potato disease leaf images. New lightweight methods such as the point-by-point convolution technique and inverse residual network structure are used to design and adjust the number of ascending and descending channel multiples, the number of convolution kernels, and the number of module repetitions, which improve the model recognition accuracy and significantly reduce the number of model parameters and computational complexity. The overfitting problem is also solved by adding a dropout layer, regularization process, and data enhancement.

Given the combined with the characteristics of the algorithms proposed in this paper, an Intel i9–9900 k CPU 3.6 GHz processor, memory of 32 GBRAM, and graphics card NVIDIA GeForce RTX 2080 Ti with 11 GB of video memory were selected for the training experimental equipment. Based on the performance characteristics of different algorithms, it is also possible to train a sufficiently good prediction model in this single GPU environment using CUDA technology.

In this paper, several deep learning convolutional neural network models have been trained using the same dataset to identify different types of potato leaf diseases. We first load the original model using Python language to build a pre-trained model without classifier. Then we add a global average pooling layer and a fully connected layer. Finally, a SoftMax function based classifier is added according to the disease leaf type.

The training process uses the weights pre-trained on ImageNet for network parameter updates. The loss function of the loss layer uses the cross-entropy loss function. The optimizer is chosen to update the weights and biases using the Adam (Adaptive Moment Estimation) optimization algorithm [27]. Adam is a method for calculating the adaptive learning rate for each parameter or weight, and it is an extension of the gradient descent optimization algorithm, gaining the advantages of both the AdaGrad and RMSProp algorithms. After tuning the hyper parameters such as learning rate of each model several times, the experimental results are statistically shown in Table 1. As can be seen from the table, the number of parameters of MobileNet series models is significantly smaller than that of residual networks and other deep learning networks, and its Top-1 accuracy is over 93%. MobileNetV2 is very similar to the original MobileNet, except that it uses inverted residual blocks with bottlenecking features. It has a drastically lower parameter count than the original MobileNet.

In this paper, the trained models are converted to H5 files. The model deployment module is built to save the trained neural network models into the Django framework. Django is an open source web application framework written in Python that adopts the framework structure pattern of MTV-Model M, View V, and Template T. While transferring the leaf images, the model also adds category labels to the images with category labels. The leaf disease category is identified and then the returned parameters are entered into the views function, the POST port number and the HOST local address of the settings function are set, and finally a static page with that image label is generated, thus realizing a potato early leaf disease identification system. We provide different reference information for software users according to different leaf types, including disease types, information on pathogen characteristics, infestation cycle, damage symptoms, control methods and management measures, etc. The pathogenic characteristics mainly describe the physiological, tissue structural, and morphological characteristics of potato leaves disturbed by pathogens or adverse environmental conditions. The infestation cycle focuses on the process of disease development in potato leaves from the beginning of one growing season to the second growing season under the action of pathogenic bacteria. The symptoms of the disease mainly describe the signs of the disease in potato leaves. Control methods mainly introduce how people should use physical and chemical control methods to effectively solve problems when potatoes encounter diseases caused by diseases, insects, and weeds.

Management measures focus on preventive measures and related drugs for agricultural and industrial control of potato diseases.

Table 1. Identification statistics of different deep learning models. The Top-1 and Top-2 accuracy refers to the performance of the model on the validation dataset. This includes activation layers, batch normalization layers, etc. Depth refers to the topological depth of the network, which is the number of layers with parameters.

Models	Size	Top-1 Accuracy	Top-2 Accuracy	Depth
Xception	88 MB	0.92	0.945	126
VGG16	528 MB	0.901	0.92	23
VGG19	549 MB	0.91	0.923	26
ResNet50	98 MB	0.937	0.943	-
ResNet101	171 MB	0.932	0.936	-
ResNet152	232 MB	0.938	0.940	-
ResNet50V2	98 MB	0.93	0.94	-
ResNet101V2	171 MB	0.941	0.95	-
ResNet152V2	232 MB	0.945	0.953	-
InceptionV3	92 MB	0.932	0.948	159
InceptionResNetV2	215 MB	0.934	0.955	572
MobileNet	16 MB	0.93	0.949	88
MobileNetV2	14 MB	0.94	0.953	88
DenseNet121	33 MB	0.93	0.936	121
DenseNet169	57 MB	0.915	0.942	169
DenseNet201	80 MB	0.92	0.938	201

4. Discussion

The incidence of crop pests and diseases tends to follow a specific pattern in terms of both location and timing, however, it also exhibits a degree of disorder, displaying variations in both space and time, including regional disparities and unpredictable, sudden, and unstable occurrences. Accurate identification and appropriate control measures for crop pests and diseases can not only ensure crop production, but also effectively reduce the amount of pesticides used and the various hazards caused by pesticides. Current pest and disease data monitoring is an important part of the agricultural monitoring system.

Research on crop pest and disease problems has mostly used image processing techniques and machine learning algorithms for effective detection and has achieved more and better results to realize effective identification and extraction of various types of crop information. However, it is very difficult to collect images of crop pests and diseases in actual production. The crop pest images collected in the field are interfered by various factors, including complex backgrounds, changing lighting conditions, and being overlapped and obscured, especially when the foreground target and background leaf color are extremely similar, which brings great difficulties to the multi-class species target recognition. Traditional machine learning algorithms cannot extract the image recognition features related to pests and diseases from the complex and variable background environment, and the established crop pest and disease recognition models are not scalable. Some recognition models can only recognize a single crop pest target, and such models can hardly meet the actual production needs. Moreover, the existing deep learning recognition models for crop pests and diseases have a huge number of parameters, and it is difficult to implement the application technology development on the mobile APP with the heavy computation and the huge model volume. Therefore, it becomes extremely important to optimize the model of convolutional neural network while improving the scale and representativeness of the number of crop pest and disease image samples.

Potatoes are annual plants, usually planted in spring and harvested in fall/autumn. We collected images of the leaves mainly in the summer. The data acquisition and preprocessing are actually the most time consuming and important. It is essential to collect image data of different potato diseases and their leaves at different periods of lesions. Due to the widespread use of pesticides, it is very difficult to collect images of different types of potato diseases at different times of the year in a sufficiently size quantity and at the same time in each plantation area. Intelligent recognition models built in such cases can interfere with the accuracy of model detection in practice, including missed detections or incorrect recognition.

The insufficient number of species that can be recognized by the model often affects the model recognition accuracy in practice. For example, old leaves will show various kinds of aging and damage, which may affect the recognition result of the model. However, in practice, farmers can use their own experience to distinguish senescent leaves from diseased leaves, and they are more concerned about the type of disease, so that they can take active and effective control measures to exclude the effect of the disease. Deep learning is very scalable. If the dataset of senescing leaves is added, the generalization ability of the model will definitely be further enhanced. Limited by our current dataset, we currently focus on identifying three types of potato leaf images for early blight, late blight, and healthy. In order to enhance the robustness of the model, random rotation, left-right flip, top-down flip, fuzzy, Gaussian noise, ray transform, and random cropping are used to increase the size and sample diversity of the data set to improve the generalization ability of the model to some extent. There are many other diseases in potatoes that also appear on the leaves, and this is the next step in our research, where we will identify more types of diseases. We show some code from our experiments on GitHub. The web link is https://github.com/whoskang/early-blight-and-late-blight-leaves, accessed on 1 January 2023. We will keep updating our code in the next research work.

Complex background changes and illumination disturbances have certain effects on the deep network recognition model. In this study, we analyze the effect of model hyper parameters such as loss function, initial weights, and learning rate of the deep network model to the model training through ablation experiments. We continuously optimize different residual structure models, convolutional operations, gradient descent, and weight optimization techniques, and explore the impact of novel activation functions including ReLU6 and h-swish on improving the nonlinear capability of recognition models and high-dimensional feature extraction. Through comparative experiments, we investigate whether different lightweight neural networks can achieve the expected goals in feature extraction of potato disease leaf images, and provide technical support to realize the rapid recognition of potato disease leaf targets in mobile APP.

5. Conclusions

In this paper, we investigate a convolutional neural network working method based on multi-scale feature fusion for early leaf identification of multiple potato diseases. We analyzed and compared multiple neural network model structures. A lightweight neural network recognition framework was constructed to significantly reduce the number of model parameters. The recognition models of three early disease images of potato leaf were trained, among which the accuracy of Top-1 recognition is over 93%. We imported the trained models into the Django framework and built a website for an early identification system of potato leaf disease to provide technical support for the implementation of a mobile-based potato leaf disease identification and early warning system.

For future work, we are interested in performing fine-grained image recognition by adopting a more advanced visual backbone, e.g., vision transformers [28]. Further, the problem may be formulated as an image retrieval task whereby advanced multi-modal retrieval techniques [29] can be of potential use. Apart from pure visual information, structure information typically in the form of graph may also be explored whereby graph matching [30], especially recent deep neural graph matching [31], can be readily used. In this regard, self-supervised learning in both vision [32] and graph structure [33] can also be studied. Last but not least, a neural architecture search [34], especially object detection level methods [35], can be explored to aid the design of more suited network architecture for the recognition task at hand. This is important to consider light-weighted models.

Author Contributions: Conceptualization, F.K. and J.L.; methodology, F.K.; software, F.W.; validation, F.K. and J.L.; formal analysis, C.W.; investigation, F.K.; resources, J.L.; data curation, F.W.; writing—original draft preparation, F.K.; writing—review and editing, J.L.; visualization, F.K.; supervision, C.W.; project administration, F.K. and C.W.; funding acquisition, J.L., F.K. and C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Natural Science Foundation of Inner Mongolia Autonomous Region of China, grant number 2021BS03038; Research Program of science and technology at Universities of Inner Mongolia Autonomous Region of China, grant number NJZY21457; National Natural Science Foundation of China, grant number 32160813; Inner Mongolia Agricultural University High-level Talent Research Start-up Project, grant number NDYB2019-27; Inner Mongolia Agricultural University High-level Talent Research Start-up Project, grant number NDYB2018-38; and National Natural Science Foundation of China, grant number 32060415.

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.

References

1. Wang, F.; Wang, C.; Song, S.; Xie, S.; Kang, F. Study on starch content detection and visualization of potato based on hyperspectral imaging. *Food Sci. Nutr.* **2021**, *9*, 4420–4430. [CrossRef]
2. Deng, C.; Xue, Y.; Liu, X.; Li, C.; Tao, D. Active transfer learning network: A unified deep joint spectral–spatial feature learning model for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 1741–1754. [CrossRef]
3. Chen, W.; Chen, J.; Zeb, A.; Yang, S.; Zhang, D. Mobile convolution neural network for the recognition of potato leaf disease images. *Multimed. Tools Appl.* **2022**, *81*, 20797–20816. [CrossRef]
4. Shan-e-Ahmed, R.; Gillian, P.; John, C.; Rajpoot, P.; Nasir, M. Automatic detection of diseased tomato plants using thermal and stereo visible light images. *PLoS ONE* **2015**, *10*, e0123262.
5. Yan-Ru, Z.; Xiaoli, L.; Ke-Qiang, Y.; Fan, C.; Yong, H. Hyperspectral imaging for determining pigment contents in cucumber leaves in response to angular leaf spot disease. *Sci. Rep.* **2016**, *6*, 27790.
6. Pandey, P.; Ge, Y.; Stoerger, V.; Schnable, J. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Front. Plant Sci.* **2017**, *8*, 1348. [CrossRef]
7. Pradhan, S.S.; Patil, R. Comparison of Deep Learning Approaches for Plant Disease Detection. In Proceedings of the International Conference on Wireless Communication, Mumbai, India, 11–12 October 2019; Springer: Singapore, 2020; pp. 575–583.
8. Wan, H.; Lu, Z.; Qi, W.; Chen, Y. Plant Disease Classification Using Deep Learning Methods. In Proceedings of the 4th International Conference on Machine Learning and Soft Computing, Haiphong City, Viet Nam, 17–19 January 2020; pp. 5–9.
9. Nazki, H.; Yoon, S.; Fuentes, A.; Park, D. Unsupervised image translation using adversarial networks for improved plant disease recognition. *Comput. Electron. Agric.* **2020**, *168*, 105117. [CrossRef]
10. Qian, W.; Yang, X.; Peng, S.; Guo, Y.; Yan, J. Learning modulated loss for rotated object detection. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–17 February 2021; Volume 35, pp. 2458–2466.
11. Brock, A.; Donahue, J.; Simonyan, K. Large scale GAN training for high fidelity natural image synthesis. *arXiv* **2018**, arXiv:1809.11096.
12. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S. SSD: Single shot multibox detector. In *European Conference on Computer Vision*; Springer: Cham, Germany, 2016; pp. 21–37.
13. 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, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
14. Yang, X.; Yan, J.; Yang, X.; Tang, J.; Liao, W.; He, T. Scrdet++: Detecting small, cluttered and rotated objects via instance-level feature denoising and rotation loss smoothing. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 2384–2399. [CrossRef]
15. 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]
16. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [CrossRef]
17. Zhong, Y.; Zhao, M. Research on deep learning in apple leaf disease recognition. *Comput. Electron. Agric.* **2020**, *168*, 105146. [CrossRef]
18. Yuan, C.; Wu, Y.; Qin, X.; Qiao, S.; Pan, Y.; Huang, P.; Liu, D.; Han, N. An effective image classification method for shallow densely connected convolution networks through squeezing and splitting techniques. *Appl. Intell.* **2019**, *49*, 3570–3586. [CrossRef]
19. Amara, J.; Bouaziz, B.; Algergawy, A. A deep learning-based approach for banana leaf diseases classification. In *Datenbanksysteme für Business, Technologie und Web (BTW 2017) – Workshopband*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2017; pp. 79–88.

20. Mohanty, S.P.; Hughes, D.P.; Marcel, S. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef]
21. Amanda, R.; Kelsee, B.; Peter, M.C.; Babuali, A.; James, L.; David, P.H. Deep learning for image-based cassava disease detection. *Front. Plant Sci.* **2017**, *8*, 1852.
22. Sladojevic, S.; Arsenovic, M.; Anderla, A.; Culibrk, D.; Stefanovic, D. Deep neural networks based recognition of plant diseases by leaf image classification. *Comput. Intell. Neurosci.* **2016**, *2016*, 3289801. [CrossRef]
23. Brahimi, M.; Boukhalfa, K.; Moussaoui, A. Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. *Appl. Artif. Intell.* **2017**, *31*, 299–315. [CrossRef]
24. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
25. Sandler, M.; Howard, A.G.; Zhu, M.; Zhmoginov, A.; Chen, L. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
26. Howard, A.; Sandler, M.; Chu, G.; Chen, L.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Salt Lake City, UT, USA, 15–20 June 2019; pp. 1314–1324.
27. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
28. Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. *ACM Comput. Surv. (CSUR)* **2022**, *54*, 1–41. [CrossRef]
29. Deng, C.; Chen, Z.; Liu, X.; Gao, X.; Tao, D. Triplet-based deep hashing network for cross-modal retrieval. *IEEE Trans. Image Process.* **2018**, *27*, 3893–3903. [CrossRef] [PubMed]
30. Yan, J.; Yin, X.; Lin, W.; Deng, C.; Zha, H.; Yang, X. A short survey of recent advances in graph matching. In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, New York, NY, USA, 6–9 June 2016; pp. 167–174.
31. Yan, J.; Yang, X.S.; Hancock, E.R. Learning for graph matching and related combinatorial optimization problems. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, Yokohama, Japan, 7–15 January 2021; pp. 4988–4996.
32. Caron, M.; Touvron, H.; Misra, I.; Jégou, H.; Mairal, J.; Bojanowski, P.; Joulin, A. Emerging properties in self-supervised vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Nashville, TN, USA, 15–20 June 2021; pp. 9650–9660.
33. Liu, C.; Zhang, S.; Yang, X.; Yan, J. Self-supervised learning of visual graph matching. In Proceedings of the European Conference on Computer Vision, New Orleans, LA, USA, 18–24 June 2022; Springer: Cham, Germany, 2022; pp. 370–388.
34. Elsken, T.; Metzen, J.H.; Hutter, F. Neural architecture search: A survey. *J. Mach. Learn. Res.* **2019**, *20*, 1997–2017.
35. Wang, X.; Lin, J.; Zhao, J.; Yang, X.; Yan, J. EAutoDet: Efficient architecture search for object detection. In *European Conference on Computer Vision*; Springer: Cham, Germany, 2022; pp. 668–684.