agriculture

MDPI

Article

Rice Plaque Detection and Identification Based on an Improved Convolutional Neural Network

Jiapeng Cui [1,2,3] and Feng Tan [2,*]

[1] College of Engineering, Heilongjiang Bayi Agricultural University, Daqing 163000, China
[2] College of Electrical and Information, Heilongjiang Bayi Agricultural University, Daqing 163000, China
[3] Branch of Suihua, Heilongjiang Academy of Agricultural Mechanization Sciences, Suihua 152054, China
* Correspondence: bayitf@byau.edu.cn

Abstract: Rice diseases are extremely harmful to rice growth, and achieving the identification and rapid classification of rice disease spots is an essential means to promote intelligent rice production. However, due to the large variety of rice diseases and the similar appearance of some rice diseases, the existing deep learning methods are less effective at classification and detection. Aiming at such problems, this paper took the spot images of five common rice diseases as the research object and constructed a rice disease data set containing 2500 images of rice bacterial blight, sheath blight, flax leaf spot, leaf streak and rice blast, including 500 images of each disease. An improved lightweight deep learning network model was proposed to realize the accurate identification of disease types and disease spots. A rice disease image classification network was designed based on the RlpNet (rice leaf plaque net) network model, Which is the underlying network, in addition to the YOLOv3 target detection network model in order to achieve the optimization of the feature extraction link, i.e., upsampling by transposed convolution and downsampling by dilated convolution. The improved YOLOv3 model was compared with traditional convolutional neural network models, including the AlexNet, GoogLeNet, VGG-16 and ResNet-34 models, for disease recognition, and the results showed that the average recall, average precision, average F1-score and overall accuracy of the network model for rice disease classification were 91.84%, 92.14%, 91.87% and 91.84%, respectively, which were all improved compared with the traditional algorithms. The improved YOLOv3 network model was compared with FSSD, Faster-RCNN, YOLOv3 and YOLOv4 for spot detection studies, and the results showed that it could achieve a mean average precision (mAP) of 86.72%, a detection rate (DR) of 93.92%, a frames per second (FPS) rate of 63.4 and a false alarm rate (FAR) of only 5.12%. In summary, the comprehensive performance of the proposed model was better than that of the traditional YOLOv3 algorithm, so this study provides a new method for rice disease identification and disease spot detection. It also had good performance in terms of the common detection and classification of multiple rice diseases, which provides some support for the common differentiation of multiple rice diseases and has some practical application value.

Keywords: rice disease; convolutional neural network; multitarget classification; real-time detection

check for updates

Citation: Cui, J.; Tan, F. Rice Plaque Detection and Identification Based on an Improved Convolutional Neural Network. *Agriculture* **2023**, *13*, 170. https://doi.org/10.3390/agriculture13010170

Academic Editor: Michele Pisante

Received: 18 November 2022
Revised: 27 December 2022
Accepted: 4 January 2023
Published: 9 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rice is one of the most important food crops in China and accounts for 30% of the country's grain output. During the growing period of rice, rice seedlings may be threatened by a variety of diseases due to the influence of many factors, such as genes and the environment. In recent years, with the influence of various factors such as climate change, rice planting systems, the selection of rice varieties and irrational drug use, the areas of rice diseases have increased significantly, accompanied by an increase in the types of diseases and aggravation of the degree of damage. The annual losses caused by rice diseases are very large; they directly threaten rice production and further threaten China's grain reserves [1,2]. Therefore, disease monitoring and control of rice is a key technical approach to rice production.

Currently, in terms of rice disease identification, the identification methods are mainly divided into manual detection, semiautomatic detection and automatic detection [3,4]. Manual detection is the visual observation of diseased spots by the human eye, upon which the disease type of rice is subjectively judged based on existing knowledge and experience. Semiautomatic detection can classify rice disease species using machine learning algorithms, such as support vector machine and random forest algorithms, provided that the basic features, such as the color, texture, and morphology of the rice disease spots, are extracted manually. Automatic detection is driven by a large amount of rice disease data using deep learning techniques with supervised or semisupervised learning, and the multilayer features of rice disease spots are automatically extracted after various nonlinear transformations to finally achieve the classification of the rice disease species. Among them, Dengshan Li et al. [5] used Fast-RCNN as a framework to detect rice diseases in video sequences, and Zhe Xu et al. [6] verified the advantage of deep convolutional networks over support vector machines (SVM) in rice leaf image classification. Song Liang et al. [7] developed a technique based on residual network (ResNet-101) for the average recognition of three rice leaf diseases, and the average accuracy of recognition of the three rice leaf diseases based on residual network (ResNet-101) was up to 99.89%.

In recent years, with the development of computer technology and agricultural informatization, many scholars have tried to use computer image processing and pattern recognition technology to achieve the automatic identification and diagnosis of crop diseases [8]. Ren Shouzang et al. [9] constructed a model based on the deconvolution-guided VGG (Visual Geometry Group) network (Deconvolution-Guided VGGNet, i.e., DGVGGNet) to achieve the identification of plant leaf disease species and disease spot segmentation and used the model to identify tomato leaf diseases. Shi-Fang Su et al. [10] proposed a grape leaf disease recognition (Grape-VGG-16, i.e., GV) model, which solved the problems of slow convergence and easy overfitting of convolutional neural networks in small-sample grape disease leaf recognition and proposed a migration-learning-based model training method for this model. Yan Guo et al. [11] could achieve an accuracy of 83.57% for the identification of three leaf diseases using region suggestion network (RPN), which is better than the traditional method. Khamparia, A. et al. [12] designed a hybrid approach for the detection of crop leaf diseases using a combination of convolutional neural networks and autoencoders. The proposed network was trained on five diseases of three crops, and it could distinguish crop diseases using leaf images and was better than other conventional methods. Jindal, U. et al. [13] proposed three versions of convolutional neural networks for the classification of plant diseases in color images, grayscale images and segmented images and experimentally verified that the proposed model had better recognition accuracy, recall and precision on color image datasets. Khalied Albarrak et al. [14] achieved the classification of eight different species of date fruits by improving MobileNet V2 and achieved a 99% classification accuracy. Yonis Gulzar et al. [15] used deep learning and migration learning to classify 14 common seed images and achieved a 99% classification accuracy on the test set.

Considering that rice diseases have different representations at different fertility stages and different growth environments, i.e., the locations of disease spots and the sizes of disease spots vary [16], the recognition accuracy in rice disease spot recognition cannot meet expectations when using the existing common convolutional neural network recognition models. The main contributions of the paper were as follows:

(1) RIpNet, which was suitable for image classification of five rice diseases, was proposed, and its classification effect was better than the existing classification network model;
(2) The YOLOv3 model was improved by using RplNet as the backbone network to make it suitable for the spot detection of five rice diseases, and the detection effect was better than the traditional YOLOv3 and YOLOv4 models;
(3) The idea of target detection was used to detect rice diseases, which could further accurately locate the occurrence area of rice diseases and provide some ideas for the accurate prevention and control of diseases.

2. Materials and Methods

2.1. Establishment of the Rice Disease Spot Image Database

2.1.1. Rice Disease Selection and Disease Spot Characteristics

The reason for the greater variety of rice diseases is that the disease spots occur at different sites during the different periods of rice growth, resulting in the presentation of different spot characteristics. Generally, the parts that are susceptible to disease are the leaves, stems and spikes [16]. According to the catalog of key diseases for the prevention and control of single-season japonica rice planting areas in northern China issued by the Technical Program for Prevention and Control of Major Rice Pests and Diseases in 2021, China Agricultural Technology Center, five common rice diseases, bacterial leaf blight, stripe blight, hoary leaf spot, stripe disease and leaf blast, were selected. Based on the characteristic descriptions of the diseases, five kinds of rice disease characteristic datasets were established, as shown in Table 1. These feature descriptions can provide a visual basis for subjective discrimination of the accuracy of rice disease image recognition by deep learning network models [17].

Table 1. Parameters of the transmission mechanism.

	Lesion Category	Lesion Features
1	White leaf blight	The location of this type of disease spot is mainly concentrated on rice leaf sheaths and leaves. The spots initially form yellow patches at the edges of the rice leaves and then expand from the leaf edges to the middle of the rice leaves, forming striated yellow–green or green–gray patches. With the further growth of the spots, the color of the spots change from yellow–green or green–gray to gray–white, while inward curling symptoms occur on the rice leaves.
2	Stripe blight	The blight initially occurs on the leaf sheaths near the water surface, and its shape shows irregular oval gray–green or light coffee-colored patches, and then the spots extend to the upper part of the rice plant, eventually forming a central gray–white patch.
3	Hoary leaf spot	The location of the disease spot is on the rice leaf sheath. As the disease intensifies, the spots form irregular oval dark coffee color patches in the center of gray–brown irregular patches. The onset of serious disease often appears as dark brown spots.
4	Stripe disease	The spots are initially dark green water-soaked spots, which soon expand into dark green spots and yellowish-brown thin stripes between the leaf veins that are approximately 1 × 10 mm in size. Because of the rapid onset of this type of disease, the poke-like patches can expand into dark green stripes in the middle of the leaf veins within a short period of time. When the disease is serious, the stripes can become connected to form a large white spot, and the leaves can become involute, resulting in a large yellow–white rice field.
5	Rice leaf blight	The spots are initially irregular dark green dots and are located on the rice leaves without a fixed location, and as the disease intensifies the dotted spots begin to expand into ovals. In severe cases, green mold is produced on the spots. The growth rate of this type of spot is extremely fast.

2.1.2. Construction of the Rice Disease Spot Image Database

A deep neural network model can recognize a large set of diseases in images, so it can achieve the rapid recognition of specified features in complex situations and then complete the recognition of images, which is similar to the extraction and learning of abstract image features from low to high levels in the human brain [18]. The lack of rice disease image information has led to the absence of a completed image library, so a large collection of rice disease spot images was necessary. In this study, one hundred valid images were collected for each of the five types of diseases using a handheld digital camera in a laboratory environment and an actual rice field environment, The camera type was Nikon D3500, the image resolution included 6000 × 4000 pixels and 4000 × 4000 pixels and the shooting time was April–October 2021. Considering that the number of samples has an impact on the recognition performance of the model, while the opposite is true for image preprocessing, the amount of original sample data was expanded 5-fold using image data enhancement, i.e., 500 images of each of the 5 diseases were obtained from the original dataset. Some of the images are shown in Figure 1. From left to right: image sharpening, histogram equalization, mirroring over the X-axis, mirroring over the Y-axis and the original image.

Furthermore, to meet the requirements of accurate disease detection, the parts of the dataset with diseases were labeled by the LabelImg software in the VOC dataset format [19,20], and then the dataset was randomly divided into training and test sets in a ratio of 3:1.

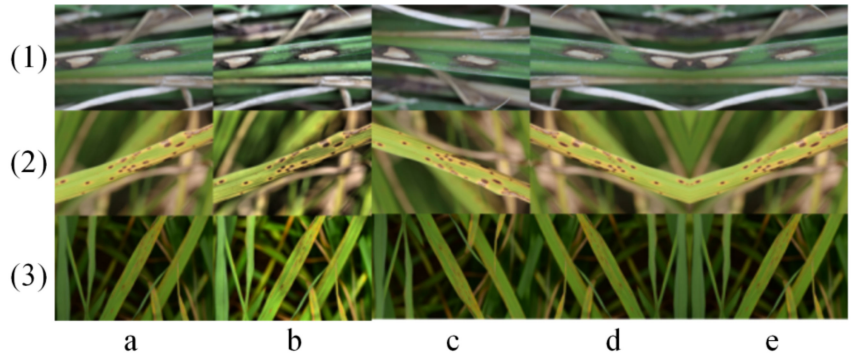


Figure 1. Disease image dataset. (**1**) is the image of stripe blight, (**2**) is the image of hoary leaf spot, (**3**) is the image of stripe disease, and (**a**–**e**) are image sharpening, histogram equalization, mirroring by X-axis, mirroring by Y-axis and original image, respectively.

2.2. Image Classification Network Model

2.2.1. Construction of the Rice Disease Spot Image Database

Due to the poor mobility of convolutional neural networks [21] and the complexity of the situations in rice fields in natural environments, it is difficult for mainstream convolutional neural network image classification models in agricultural applications to guarantee both classification accuracy and economic efficiency, leading to the inability to directly select these models for the work in our study. To construct a network model (RlpNet) applicable to the automatic classification of the above five common rice disease images, the LeNet, AlexNet, GoogLeNet, VGG and ResNet network models were referenced and combined [22–25]. The model mainly consists of interspersed continuous convolution and Inception structures. In addition, considering that the recent mainstream residual networks mainly address the problem of explosion gradients in deep network models [26], while the network model proposed in this study has fewer layers, a residual network module was not added to the network model.

2.2.2. Network Model Framework

The structure of the image classification network model is shown in Figure 2. The network model contains a 30-layer network structure, with layer 1 being the input layer and layer 2 being the convolutional layer, which has 32 channels, a convolutional kernel size of 3 × 3 and a step size of 1. It also contains a batch normalization (BN) [27] layer and a ReLU activation function [28], where the BN approach is a method that was proposed by Google for optimizing convolutional neural network training and has some applications in GoogLeNet's Inception V2 model. The third layer is a convolutional layer that is similar to the second layer, but the number of channels is increased from 32 to 64 and the convolutional step size is 2. The purpose is to reduce the dimensionality of the image instead of maximum pooling and to prevent the disappearance of features during maximum pooling as much as possible. Layer 4 is the maximum pooling layer with a pooling window size of 2 and a step size of 1 and is mainly used to extract salient features in the image. Layers 4 to 12 are inception structures, which are used to extend the width of the network model and improve the operational efficiency while making the image feature extraction more adequate. Layers 13 to 16 are a combination of two maximum pooling layers and two convolutional layers for further reducing image dimensionality and extracting image features in different dimensions. Layers 16 to 24 have the same inception structure as above and are placed before the output structure to further improve the operational efficiency and deepen the multifeature extraction task. Next is a convolutional layer with a convolutional kernel size of 1 × 1 and a maximum pooling layer with a pooling window of 3 × 3. On

the one hand, this layer is used to change the channel dimension to ensure the information interaction between different channels and increase the nonlinear characteristics of the model, and on the other hand, features at different scales can be extracted to prevent the occurrence of the overfitting phenomenon. Layer 27 is a fully connected layer. Layer 28 is a dropout layer, which can improve the generalization ability of the network model, solve the overfitting problem of the network model to a certain extent and optimize the training of the network model. Layers 29 and 30 are the softmax activation function layer and the classification output layer, respectively.

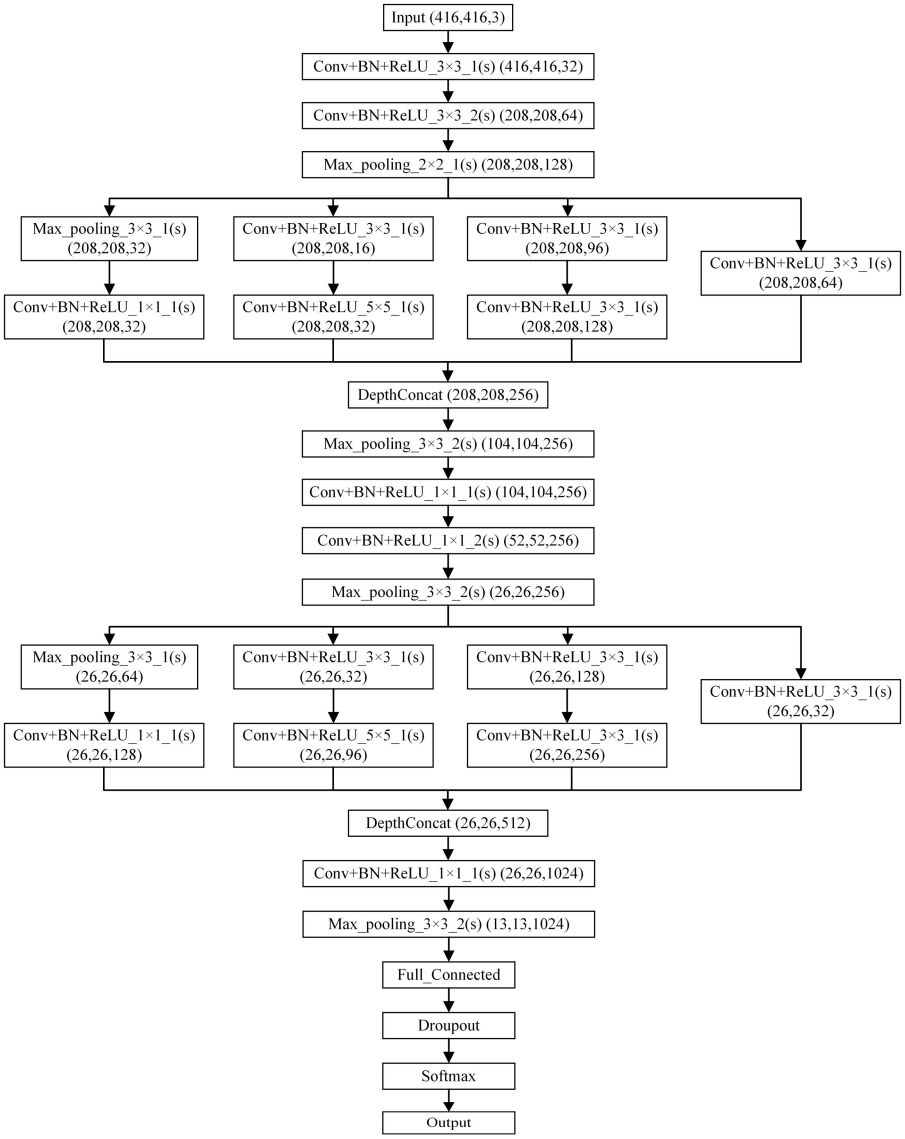


Figure 2. Lesion classification model map.

In the model structure part of the operation, the image is downscaled using stride convolution [29] instead of maximum pooling. The principle is shown in Figure 3.

As seen in Figure 3, the stride convolution operation is similar to ordinary convolution, but the scan of the image by the convolution kernel skips some regions, thus achieving image dimensionality reduction in the final output. In contrast, max pooling is the process of traversing an image in the pooling window and retaining only the most prominent feature information in the pooling window, thus enabling the output image to be dimensionally reduced. From the comparison, it can be concluded that these are two different methods of image dimensionality reduction and feature extraction, and both operations were applied

in the model of this study to ensure the diversity and richness of feature extraction in the feature model.

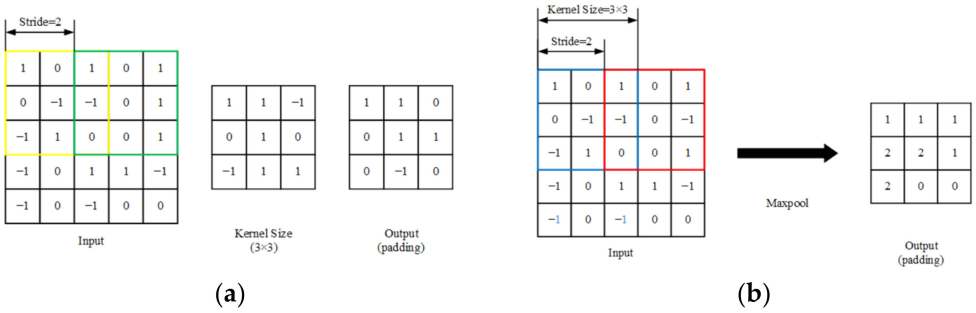


Figure 3. Principle of dimensionality reduction. (**a**) Stride convolution; (**b**) max pooling.

2.3. Disease Target Detection Network Model

In the process of target detection of rice diseases, different diseases have different characteristics and exhibit polymorphic forms, and the morphology of the same disease spot varies at different periods [30,31]. Disease spots have certain growth characteristics from generation to outbreak, so in the process of detection, the model needs to have good recognition of both obvious and tiny targets. Therefore, it is necessary to extract deeper network features during feature extraction.

To ensure the final target detection accuracy and efficiency, the mentioned image classification network is used as the backbone network of the disease target detection network model; the final fully connected layer, the dropout layer, the softmax layer and the output layer are removed from the YOLOv3 network model [32,33]; the output features are extracted at three different scales in layers 15, 25 and 26; and feature fusion by means of up- and downsampling is applied. The target detection network model is shown in Figure 4. The number on the convolution block in the figure represents the number of such convolution blocks; otherwise, the number of convolution blocks is 1.

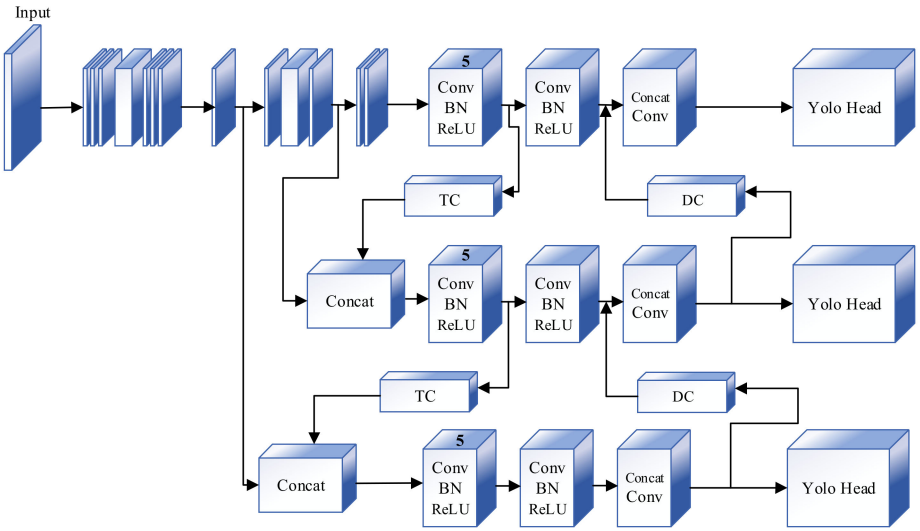


Figure 4. Disease target detection model diagram.

In the feature extraction, the upsampling module of YOLOv3 is used to ensure that the image features can be fully extracted and fused. However, to ensure that the original image features are maintained during the sampling process to improve the recognition of fine-grained images, the bilinear upsampling operation in the upsampling process is changed to transposed convolution (TC) [34,35]. In addition, we added a subsampling operation between the three feature output channels. During downsampling, the perceptual field is

expanded by dilated convolution (DC) to increase multiscale information [36]. The model features are extracted by TC and DC to increase the fusion of multiscale features, thus improving the detection capability of the model and the recognition effect on obvious and tiny targets.

TC, also known as deconvolution, can be understood as the inverse process of convolution. Its operation is shown in Figure 5. The main purpose of TC is to highlight small features in an image by reducing the underlying image to the upper layer or to a larger image size through convolution operations. Compared with upsampling, the TC operation can tap into each pixel feature and restore finer local features, which is more suitable for detection work on small targets. However, deconvolution is prone to uneven overlap, especially when the kernel size is not divisible by the step size, resulting a checkerboard pattern. Therefore, the TC kernel size is set to 2×2, the convolution step size is set to 1 to alleviate this situation and activation is performed with the ReLU function after convolution.

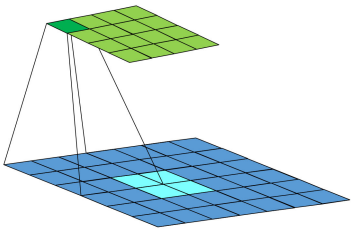


Figure 5. Transposed convolution operation.

Dilated convolution, also called inflated convolution, is based on the standard convolution operation by filling the middle of the convolution kernel with zeros to expand the field of perception in a deep network model and capture multiscale information. In target detection, the more multiscale information that is fused to a certain extent, the higher the accuracy of target detection. The difference between dilated convolution and convolution is shown in Figure 6.

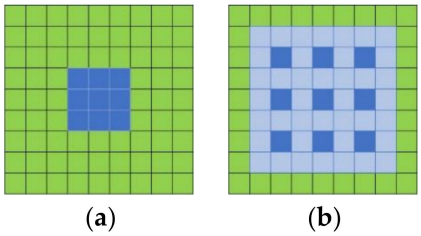


Figure 6. Convolution process. (**a**) Convolution; (**b**) dilated convolution.

Since dilated convolution fills the convolution kernel with zeros, it theoretically increases the size of the convolution kernel. However, since it does not introduce parameters of practical significance, it does not increase the computational effort. The size of the convolution kernel in dilated convolution is calculated by the following equation.

$$K = k + (k - 1) \times (r - 1) \quad (1)$$

where K is the size of the convolution kernel, k is the size of the original convolution kernel and r is the dilated rate.

Dilated convolution has the property that it cannot appear in convolution operations where the convolution kernel is 1×1. However, the general convolution kernel size is not 2×2; therefore, to reduce the number of parameters, a convolution kernel of size 3×3 is used in the dilated convolution operation introduced in this study, and the dilated rate was chosen to be 2.

3. Results

The image classification model and the target detection model were trained separately using the rice disease database constructed in Section 1. The hardware configuration used in the experiments was an Intel (R) Core (TM) i7-8550 CPU and an NVIDIA GeForce RTX 2080 super graphics card. The operating system was Windows 10 OS, and the deep learning environment was configured with PyTorch 1.6.1 + CUDA 10.1 + CUDNN 7.6.5 + Python 3.6.9.

3.1. Experimental Results of Disease Image Classification

In the classification experiments, the training set data were fed into the model for training. AlexNet, VGG-11, GoogLeNet and ResNet-34 were selected for comparison [37]. The training period was set to 1500, the sample batch size was set to 16, the initial learning rate was set to 0.01 and the learning rate was gradually reduced by using the learning rate step-by-step approach. The trend of correctness of each network model with an increasing iteration period is shown in Figure 7.

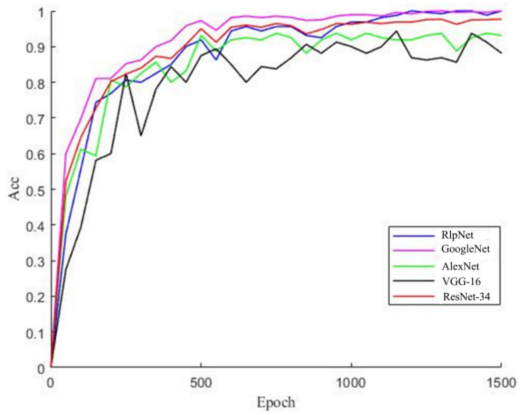


Figure 7. Training accuracy of the different networks.

From Figure 7, it can be concluded that the training accuracy of all five types of models increased rapidly with the increase in the training time during the first 200 training cycles and stabilized with small fluctuations after 500 training cycles. The training accuracy of the network model proposed in this study and GoogLeNet converged to 1 after 1200 training cycles. ResNet-34 eventually stabilized at approximately 0.97. In contrast, VGG-16 and AlexNet performed poorly in training and fluctuated more significantly after stabilization. This was caused by the shallower depth and simpler structure of these two types of models compared to those of the other network models, which led to a poorer learning ability for multiscale features.

After the network model was trained, the network model was tested using the rice disease spot test set images. In the test, the model parameters with the best training effect for each network model were called. The following test evaluation metrics were selected to evaluate the models: accuracy, recall, precision and F1 score. The calculation formulas are as follows:

$$\mathrm{Ac} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP} + \mathrm{TN}} \times 10, \tag{2}$$

$$\mathrm{Re} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100\% \tag{3}$$

$$\mathrm{Pr} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\% \tag{4}$$

$$\text{F1-score} = 2 \times \frac{\mathrm{Re} \times \mathrm{Pr}}{\mathrm{Re} + \mathrm{Pr}} \tag{5}$$

where Ac denotes accuracy, Re denotes recall, Pr denotes recognition accuracy, TP denotes the number of true positive samples, TN denotes the number of true negative samples,

FN denotes the number of false negative samples and FP denotes the number of false positive samples. In this study, taking rice blight as an example, for the blight images, the number of blight images in the sample that are correctly identified as blight is denoted as TP; the number of other disease images in the sample data that are correctly identified as the corresponding disease is denoted as TN; the number of blight images in the dataset that are identified as other diseases is denoted as FN; and the number of images of other diseases in the dataset that are identified as blight is FP.

The aforementioned five image classification models were tested on the test set images, and the final results were evaluated with the above four evaluation indexes. The final test results are shown in Table 2.

Table 2. Test results.

		Bacterial Leaf Blight	Stripe Blight	Hoary Leaf Spot	Stripe Disease	Leaf Blast	Average	Accuracy	Time (s)
Goog-LeNet	Recall	84.80%	78.40%	92.80%	89.60%	86.40%	86.40%	86.40%	2.6413
	Accuracy	**99.07%**	79.03%	75.82%	94.12%	88.50%	87.31%		
	F1 score	91.38%	78.71%	83.45%	91.80%	87.45%	86.56%		
AlexNet	Recall	68.00%	76.80%	80.80%	79.20%	85.60%	78.08%	78.08%	3.0842
	Accuracy	77.98%	78.05%	68.71%	76.74%	91.45%	78.59%		
	F1 score	72.65%	77.42%	74.26%	77.95%	88.43%	78.14%		
VGG-16	Recall	80.80%	72.80%	86.40%	84.00%	81.60%	81.12%	81.12%	2.7816
	Accuracy	87.83%	79.13%	81.82%	80.77%	76.70%	81.25%		
	F1 score	84.17%	75.83%	84.05%	82.35%	79.07%	81.09%		
ResNet-34	Recall	88.00%	**88.80%**	92.80%	85.60%	84.00%	87.84%	88.32%	**2.0462**
	Accuracy	97.35%	86.72%	**87.22%**	89.17%	80.15%	88.12%		
	F1 score	92.44%	87.75%	89.92%	87.35%	82.03%	87.90%		
RlpNet	Recall	**92.80%**	87.20%	**98.40%**	**91.20%**	**89.60%**	**91.84%**	**91.84%**	2.1846
	Accuracy	97.50%	**88.60%**	85.40%	**97.40%**	**91.80%**	**92.14%**		
	F1 score	**95.08%**	**87.90%**	**91.45%**	**94.21%**	**90.69%**	**91.87%**		

Note: the values marked in the table are best identified.

The results in Table 2 show that the image classification network model proposed in this study had more obvious advantages for the recognition and classification of the five types of spots in most images, and its average recall, average precision, average F1 score and overall accuracy were 91.84%, 92.14%, 91.87% and 91.84%, respectively. Only the classification accuracy for white leaf blight was slightly lower than that of GoogLeNet; moreover, the recognition classification recall for stripe blight, the classification accuracy for hoary leaf spot disease and the overall test time were slightly lower than those of ResNet-34. Overall, the network model proposed in this study had more obvious advantages in targeting the problem of rice disease classification.

The confusion matrix of RlpNet network model proposed in this study for each type of disease tested is shown in Table 3. The confusion matrix of GoogLeNet network model for each type of disease tested is shown in Table 4. The confusion matrix of the ResNet-34 network model for each type of disease tested is shown in Table 5.

Table 3. Confusion matrix of RlpNet network model.

	Bacterial Leaf Blight	Stripe Blight	Hoary Leaf Spot	Stripe Disease	Leaf Blast	Total
Bacterial leaf blight	116	3	2	3	1	125
Stripe blight	0	109	13	0	3	125
Hoary leaf spot	0	1	123	0	1	125
Stripe disease	3	3	0	114	5	125
Rice leaf blast	0	7	6	0	112	125
Total	119	123	144	117	122	625

Table 4. Confusion matrix of GoogLeNet network model.

	Bacterial Leaf Blight	Stripe Blight	Hoary Leaf Spot	Stripe Disease	Leaf Blast	Total
Bacterial leaf blight	106	7	8	2	2	125
Stripe blight	0	98	23	1	3	125
Hoary leaf spot	0	1	116	4	4	125
Stripe disease	1	5	2	112	5	125
Rice leaf blast	0	13	4	0	108	125
Total	107	124	153	119	122	625

Table 5. Confusion matrix of ResNet-34 network model.

	Bacterial Leaf Blight	Stripe Blight	Hoary Leaf Spot	Stripe Disease	Leaf Blast	Total
Bacterial leaf blight	110	5	2	2	6	125
Stripe blight	0	111	4	3	7	125
Hoary leaf spot	1	2	116	4	1	125
Stripe disease	2	3	1	107	12	125
Rice leaf blast	0	6	10	4	105	125
Total	113	128	133	120	131	625

3.2. Experimental Results of Lesion Detection

The labeled dataset in Section 3.1 above was used for training and testing, and the FSSD, Faster-RCNN, YOLOv3 and YOLOv4 algorithms were selected for comparison. In the experiments, the model was trained by freeze training. The number of freeze training iterations and nonfreeze training iterations were set to 200 and 300, respectively. The sample batch size was set to 8, and the learning rate was set to the same as that of the image classification network. In the training for each network model, to ensure the consistency of the experimental results, the network models did not use pretraining weights, and all models started training uniformly again.

After training, the training weights corresponding to the lower loss values of each target detection network model were called for testing. Therefore, in addition to the common mean average precision (mAP) and frames per second (FPS) evaluation indexes, the false alarm rate (FAR) and detection rate (DR) indexes were added to evaluate the detection effectiveness of the network model. The calculation of each evaluation index is as follows:

$$\mathrm{mAP} = \frac{\sum_{i=1}^{n} \mathrm{AP(i)}}{n} \tag{6}$$

where n is the number of categories classified in the dataset used, and AP is the average detection accuracy for each category at different recall rates.

$$\mathrm{FAR} = \frac{\mathrm{APi}}{\mathrm{TPi}} \tag{7}$$

$$\mathrm{DR} = \frac{\mathrm{DPi}}{\mathrm{TPi}} \tag{8}$$

where APi is the number of test set images with false alarm detection results, which mainly include detection errors. Dpi is the number of images with all spots detected in the image. Tpi is the total number of test set images.

From Table 6, it can be concluded that for the detection problem of the five kinds of rice disease spots, the mAP of the improved algorithm in this study could reach 86.72%, which was better than that of the traditional YOLOv3 algorithm, significantly better than those of FSSD and Faster-RCNN and was even better than that of the YOLOv4 algorithm. The innovative algorithm in this study also had a more obvious advantage in terms of the FAR value, which was only 5.12%. The DR was 93.92%, and the FPS was 63.4, which was slightly

inadequate compared to the YOLOv4 algorithm. This was because this target detection algorithm uses TC and DC to achieve upsampling and downsampling, respectively, in the feature fusion part, and hence fuses more scales of feature information, which ensures better detection accuracy and lower FAR values. However, since this target detection algorithm uses a shallower backbone network and does not use a residual module, there may be some missed detections, resulting in a slightly lower DR. Weighing the evaluation indexes of mAP, FAR, DR and FPS4, the innovative network model in this study still had certain advantages to meet the demand for the real-time detection of rice seedling disease spots in a field environment.

Table 6. Detection effects of the different algorithms.

Models	mAP/%	DR/%	FAR/%	FPS
FSSD	75.36	88.8	10.2	41.4
Faster-RCNN	70.56	89.44	12.5	21.1
YOLOv3	81.76	94.56	8.17	61.2
YOLOv4	84.48	**95.2**	5.31	**65.8**
Innovative YOLOv3	**86.72**	93.92	**5.12**	63.4

Note: the values marked in the table are best identified.

3.3. Comparison of Disease Image Classification and Disease Spot Detection

From the two sets of experimental results in Sections 3.1 and 3.2, it was concluded that the classification accuracy of the image classification network proposed in this study could reach more than 91% for the five types of rice disease images, while the mAP for detecting the five types of disease spots was only 86.72%, which means the effect of image classification was significantly better than the effect of disease spot target detection in the images. The spot detection test results are shown in Figure 8. The reason for the difference was that the two types of algorithms act on different objects. The image classification network only needs to determine the kind of disease spots contained in the image to classify them as the corresponding disease spots. This classification model can extract more image information that does not need to be precisely located, including the information of the leaves where the disease spots grow and other information of the image in addition to the disease spots, which prevents the interference of the background in the image. Spot target detection is equivalent to further spot classification, which requires precise marking of the location of a spot in an image based on the classification of the spot and is therefore subject to greater interference and errors caused by mislabeling of the spot location, missed marking of the spot, etc. In addition, different forms of the same disease spot or a different sharpness of the disease spot may exist in an image and may interfere with disease spot detection. Therefore, the target detection algorithm was slightly less accurate than the corresponding image classification algorithm. However, since the target detection algorithm can mark more accurate location information, it has greater application advantages in scenes that require accurate detection or when an image contains multiple targets.

At the early stage of a disease outbreak, none of the disease characteristics are obvious, and it is difficult to capture disease information by means of wide-scale detection, such as drones. While using the rice disease spot image detection and identification method in this study, the disease spot information contained in a single image could be artificially controlled by means of manual photography. In addition, due to the complexity of the paddy field environment and the limitations of agricultural machinery and equipment, there is no regularity in disease distribution when multiple diseases break out in a rice field at the same time, resulting in pesticide-spraying challenges for control and precise-spraying challenges for a single disease. In such cases, ensuring the classification accuracy of rice disease spot images means that the initial judgment of the disease extent can be satisfactory and can provide a basis for wide-scale disease control. Precise disease target detection will improve with more precise agronomic and farm machinery facilities.

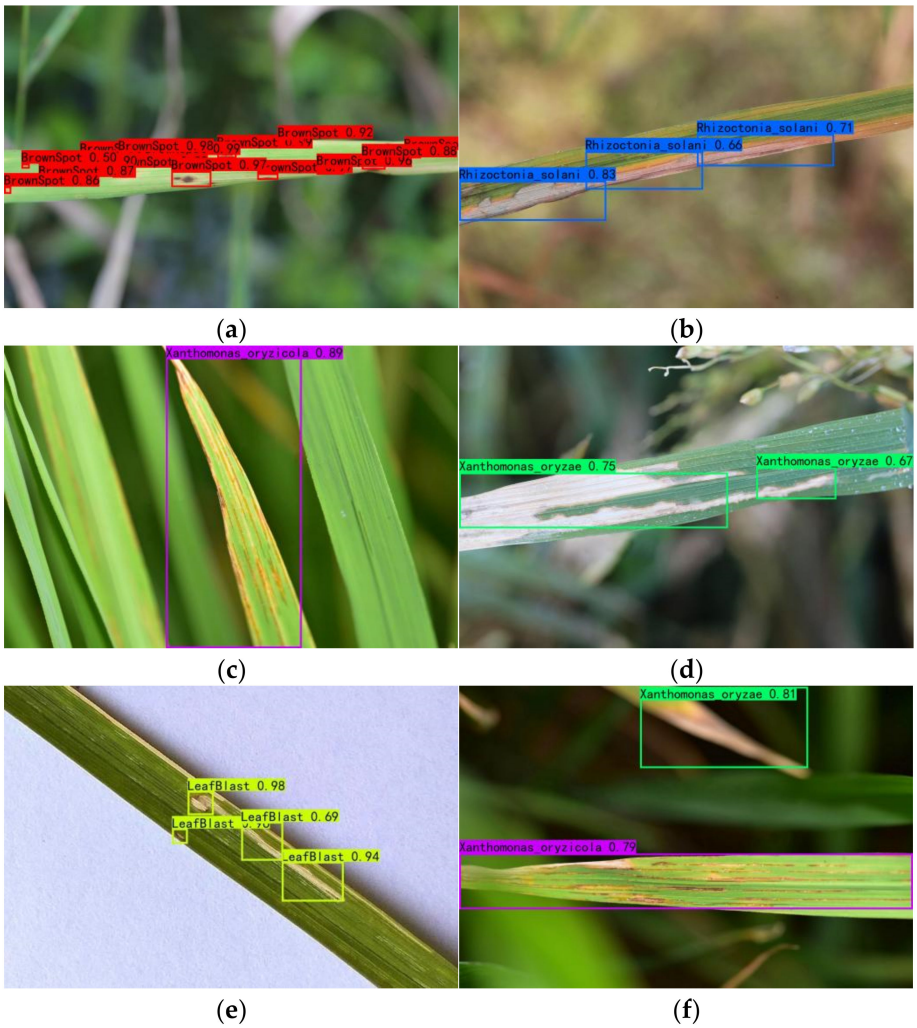


Figure 8. Spot detection test results. (**a**) In the red box is hoary leaf spot; (**b**) in the blue box is stripe blight; (**c**,**f**) in the purple box is stripe disease; (**d**,**f**) in the green box is white leaf blight; (**e**) in the green box is rice leaf blight.

4. Discussion

Deep learning has many applications in agriculture, including disease classification and detection, seed classification, fruit classification and defect detection, etc. In addition, some deep learning methods have proved their effectiveness in corresponding agricultural applications [8–15]. In this paper, the identification and detection of white leaf blight, stripe blight, hoary leaf spot, stripe disease and leaf rice blight were studied.

The core of disease classification using deep learning lies in fully extracting image features containing diseases, which mainly relies on the network model of image classification. Among them, Khalied Albarrak et al. [14] verified the effectiveness of AlexNet, VGG-16, GoogLeNet and ResNet in jujube image classification. On this basis, combined with the advantages of the above classification models, a new RIpNet was constructed to classify rice disease images. The experimental results showed that the overall classification accuracy, average classification accuracy, average recall rate and average F1 value of RlpNet for the five disease images were 91.84%, 92.14%, 91.84% and 91.87%, which were obviously improved compared with AlexNet, VGG-16, GoogLeNet and ResNet, and the maximum improvement in the overall classification accuracy, average classification accuracy, average recall rate and average F1 value was 13.76%, 13.55%, 13.76% and 13.73%, respectively. This was because RIpNet combines an inception structure on the basis of continuous convolution, and strided convolution and max pooling in the convolution process, which not only

ensures the rapid convergence of the model, but also highlights the prominent features in the image as much as possible, thus improving the performance and generalization ability of the model.

In the process of rice disease spot detection, the detection and marking of the disease spot area is the key, which mainly relies on the feature extraction of the backbone network and the feature fusion of the target detection network. Dengshan Li et al. [5] verified Faster-RCNN and different backbone networks in a study on the reliability of YOLOv3 and other target detection models of the network in the detection of three kinds of rice disease spots. The RlpNet described above in this article is the backbone network. The network improves the YOLOv3 algorithm and replaces the traditional up and down sampling structure with dilated convolution and transposed convolution in the feature fusion part, so as to better realize the fusion of subtle features and prominent features, thus realizing the accurate detection of multiple rice disease spots. The experimental results showed that the average accuracy of the improved YOLOv3 model for detecting the five rice disease spots was 86.72%, which was 16.16% and 4.96% higher than that of Faster-RCNN and YOLOv3, respectively, and the false alarm rate was only 5.12%, which was 7.38% and 3.05% higher than that of Faster-RCNN and YOLOv3, respectively, thanks to the full extraction of the disease spot features from the backbone network and the fusion of multi-scale features in the feature fusion part. It was also due to excessive multi-dimensional feature fusion that the detection rate of the improved YOLOv3 was slightly lower than that of the traditional YOLOv3, which was only 93.92%.

5. Conclusions

The research showed that it is feasible to classify and detect common rice diseases by using deep learning methods, and the method developed in this study had a high precision, high detection speed and high robustness. However, due to the complexity and diversity of rice leaf diseases in the natural environment, it is usually difficult for a deep-learning-based detection model to meet the detection requirements for all rice diseases. This paper focused on the identification and detection of five common rice diseases, including white leaf blight, stripe blight, hoary leaf spot, stripe disease and rice leaf blight.

In the aspect of disease identification research, a new target detection model, the backbone network RlpNet, was constructed from the existing deep learning detection models. First, the model rapidly reduces the size of the feature image with two consecutive convolutions, thus improving the processing speed of the model. In the model, the overall width of the network model is improved with two tandem inception structures, which has the effect of generalizing the model and further improving the processing and running speed of the model, realizing the fast dimensionality reduction in features by continuous convolution and maximum pooling between Inception, and extracting the prominent features in the image, so as to achieve more in-depth feature extraction of multiple rice diseases, which showed that the trunk network had obvious advantages in the classification accuracy of the five rice diseases.

In the research of disease spot detection, the YOLOv3 algorithm was improved by using a high-performance classification model as the backbone network, and the dilated convolution and transposed convolution are used in the feature fusion part instead of the traditional up and down sampling structure to better achieve the fusion of subtle features and prominent features, so as to achieve the accurate detection of multiple rice diseases. The experimental results showed that the disease image classification network model and the disease spot detection algorithm had certain advantages; it could quickly and accurately classify, recognize and detect five kinds of rice disease images. In addition, a small image dataset of five common rice diseases, such as white leaf blight, stripe blight, hoary leaf spot, stripe disease and rice leaf blight, was preliminarily constructed in the study.

Although the methods proposed in this paper had good effects on the identification and detection of white leaf blight, stripe blight, hoary leaf spot, stripe disease and rice leaf blight, they are still only a small part of a wide variety of rice diseases. In future work, we

need to continue to verify the applicability of this method in the identification and detection of other rice diseases and further consider the possibility of deploying this method on mobile terminals, so as to improve the practicability of this method.

Author Contributions: J.C. and F.T. conceived the study and designed the project. J.C. performed the experiment, analyzed the data and drafted the manuscript. F.T. helped to revise the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Natural Science Fund Key Project of Heilongjiang Province (ZD2019F002) and School Startup Plan (XDB2013-18).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

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

References

1. Xu, J.; Wang, J.; Xu, X.; Ju, S.C. Image recognition for different developmental stages of rice by RAdam deep convolutional neural networks. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2021**, *37*, 143–150. (In Chinese with English Abstract)
2. Tan, Y.-L.; Ouyang, C.-J.; Li, L.; Pengjie, T.T. Image recognition of rice diseases based on deep convolutional neural network. *J. Jinggangshan Univ. (Nat. Sci.) Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2019**, *40*, 31–38. (In Chinese with English Abstract)
3. Turkoglu, M.; Hanbay, D. Plant disease and pest detection using deep learning-based features. *Turk. J. Electr. Eng. Comput. Sci.* **2019**, *27*, 1636–1651. [CrossRef]
4. Zhai, Z.; Cao, Y.; Xu, H.; Yuan, P.; Wang, H. Review of key techniques for crop disease and Pest detection. *Trans. Chin. Soc. Agric. Mach.* **2021**, *52*, 1–18.
5. Li, D.; Wang, R.; Xie, C.; Liu, L.; Zhang, J.; Li, R.; Wang, F.; Zhou, M.; Liu, W. A Recognition Method for Rice Plant Diseases and Pests Video Detection Based on Deep Convolutional Neural Network. *Sensors* **2020**, *20*, 578. [CrossRef]
6. Xu, Z.; Guo, X.; Zhu, A.; He, X.; Zhao, X.; Han, Y.; Subedi, R. Using Deep Convolutional Neural Networks for Image-Based Diagnosis of Nutrient Deficiencies in Rice. *Comput. Intell. Neurosci.* **2020**, *2020*, 7307252. [CrossRef]
7. Liang, S.; Deng, X. Research on rice leaf disease identification based on ResNet. In Proceedings of the IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Beijing, China, 3–5 October 2022; IEEE: Piscataway, NJ, USA, 2022.
8. Zhang, J.; Wang, H.; Guo, Y.; Hu, X. Review of deep learning. *Appl. Res. Comput.* **2018**, *35*, 1921–1928;1936.
9. Ren, S.; Jia, F.; Gu, X.; Yuan, P.; Xie, W.; Xu, H. Recognition and segmentation model of tomato leaf diseases based on deconvolution-guiding. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2020**, *36*, 186–195. (In Chinese with English Abstract)
10. Su, S.; Qiao, Y.; Rao, Y. Recognition of grape leaf diseases and mobile application based on transfer learning. *Trans. Chin. Soc. Agric. Eng. (Trans. CSAE)* **2021**, *37*, 127–134. (In Chinese with English Abstract)
11. Guo, Y.; Zhang, J.; Yin, C.; Hu, X.; Zou, Y.; Xue, Z.; Wang, W. Plant disease identification based on deep learning algorithm in smart farming. *Discret. Dyn. Nat. Soc.* **2020**, *2020*, 2479172. [CrossRef]
12. Khamparia, A.; Saini, G.; Gupta, D.; Khanna, A.; Tiwari, S.; de Albuquerque, V.H.C. Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network. *Circuits Syst. Signal Process.* **2020**, *39*, 818–836. [CrossRef]
13. Jindal, U.; Gupta, S. Deep Learning-Based Knowledge Extraction from Diseased and Healthy Edible Plant Leaves. *Int. J. Inf. Syst. Model. Des.* **2021**, *12*, 67–81. [CrossRef]
14. Albarrak, K.; Gulzar, Y.; Hamid, Y.; Mehmood, A.; Soomro, A.B. A Deep Learning-Based Model for Date Fruit Classification. *Sustainability* **2022**, *14*, 6339. [CrossRef]
15. Gulzar, Y.; Hamid, Y.; Soomro, A.B.; Alwan, A.A.; Journaux, L. A convolution neural network-based seed classification system. *Symmetry* **2020**, *12*, 2018. [CrossRef]
16. Qi, L.; Zhang, T.; Zeng, J.; Li, C.; Li, T.; Zhao, Y.; Yan, S. Analysis of the occurrence and control of diseases in five major rice-producing areas in China in recent years. *China Plant Prot.* **2021**, *41*, 37–42+65.
17. Fu, Q.; Huang, S. *Diagnosis and Control of Rice Diseases and Insect Pests*; Machinery Industry Press: Beijing, China, 2019; pp. 2–83.
18. 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**. [CrossRef]
19. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–9.
20. Fan, W.; Liu, T.; Huang, R.; Guo, Q.; Zhang, B. Convolutional neural network low-level feature-assisted image instance segmentation method. *Comput. Sci.* **2020**, *47*, 186–191.
21. Zhuang, F.Z.; Luo, P.; He, Q.; Shi, Z.Z. Survey on transfer learning research. *Ruan Jian Xue Bao/J. Softw.* **2015**, *26*, 26–39. (In Chinese)

22. Singh, V.; Misra, A.K. Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf. Process. Agric.* **2017**, *4*, 41–49. [CrossRef]
23. Seydi, S.T.; Hasanlou, M. A New Structure for Binary and Multiple Hyperspectral Change Detection Based on Spectral Unmixing and Convolutional Neural Network. *Measurement* **2021**, *186*, 11037. [CrossRef]
24. Rani, R.; Singh, G.B.; Sharma, N.; Kakkar, D. Identification of Tomato Leaf Diseases Using Deep Convolutional Neural Networks. *Int. J. Agric. Environ. Inf. Syst. (IJAEIS)* **2021**, *12*, 1–22.
25. Dat, T.T.; Le Thien Vu, P.C.; Truong, N.N.; Anh Dang, L.T.; Thanh Sang, V.N.; Bao, P.T. Leaf Recognition Based on Joint Learning Multiloss of Multimodel Convolutional Neural Networks: A Testing for Vietnamese Herb. *Comput. Intell. Neurosci.* **2021**, *2021*, 5032359. [CrossRef] [PubMed]
26. Zhuang, J.; Li, X.; Bagavathiannan, M.; Jin, X.; Yang, J.; Meng, W.; Li, T.; Li, L.; Wang, Y.; Chen, Y.; et al. Evaluation of different deep convolutional neural networks for detection of broadleaf weed seedlings in wheat. *Pest Manag. Sci.* **2021**, *78*, 521–529. [CrossRef] [PubMed]
27. De Bem, P.P.; de Carvalho Júnior, O.A.; de Carvalho, O.L.F.; Gomes, R.A.T.; Guimarães, R.F.; Pimentel, C.M.M. Irrigated rice crop identification in Southern Brazil using convolutional neural networks and Sentinel-1 time series. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100627.
28. Zhang, P.; Yin, Z.; Chen, W.; Jin, Y. CNN-Based Intelligent Method for Identifying GSD of Granular Soils. *Int. J. Geomech.* **2021**, *21*, 04021229. [CrossRef]
29. Lo Giudice, A.; Ronsivalle, V.; Spampinato, C.; Leonardi, R. Fully Automatic Segmentation of The Mandible Based on Convolutional Neural Networks (CNNs). *Orthod. Craniofacial Res.* **2021**, *24*, 100–107. [CrossRef]
30. Hu, J.; Zou, Y.; Sun, B.; Yu, X.; Shang, Z.; Huang, J.; Jin, S.; Liang, P. Raman spectrum classification based on transfer learning by a convolutional neural network: Application to pesticide detection. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *265*, 120366. [CrossRef]
31. Su, Z.; Liu, H.; Qian, J.; Zhang, Z.; Zhang, L. Hand Gesture Recognition Based on sEMG Signal and Convolutional Neural Network. *Int. J. Pattern Recognit. Artif. Intell.* **2021**, *35*, 2151012. [CrossRef]
32. Yang, B.; Zhang, Z.; Yang, C.; Wang, Y.; Orr, M.C.; Wang, H.; Zhang, B. Identification of Species by Combining Molecular and Morphological Data Using Convolutional Neural Networks. *Syst. Biol.* **2021**, *71*, 690–705. [CrossRef]
33. Zhang, D.; Li, P.; Zhao, L.; Xu, D.; Lu, D. Texture compensation with multi-scale dilated residual blocks for image denoising. *Neural Comput. Appl.* **2021**, *33*, 12957–12971. [CrossRef]
34. Maurya, R.; Pathak, V.K.; Burget, R.; Dutta, M.K. Automated detection of bioimages using novel deep feature fusion algorithm and effective high-dimensional feature selection approach. *Comput. Biol. Med.* **2021**, *137*, 104862. [CrossRef]
35. Moon, T.; Park, J.; Son, J.E. Prediction of the fruit development stage of sweet pepper (*Capsicum annum* var. *annuum*) by an ensemble model of convolutional and multilayer perceptron. *Biosyst. Eng.* **2021**, *210*, 171–180. [CrossRef]
36. Zheng, Z.; Xiong, J.; Lin, H.; Han, Y.; Sun, B.; Xie, Z.; Yang, Z.; Wang, C. A Method of Green Citrus Detection in Natural Environments Using a Deep Convolutional Neural Network. *Front. Plant Sci.* **2021**, *12*, 705737. [CrossRef]
37. Jodas, D.S.; Yojo, T.; Brazolin, S.; Velasco, G.D.; Papa, J.P. Detection of Trees on Street-View Images Using a Convolutional Neural Network. *Int. J. Neural Syst.* **2021**, *32*, 2150042. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.