

Article

Maize Small Leaf Spot Classification Based on Improved Deep Convolutional Neural Networks with a Multi-Scale Attention Mechanism

Chenghai Yin ^{1,2}, Tiwei Zeng ³, Huiming Zhang ⁴, Wei Fu ⁴, Lei Wang ^{1,2,*} and Siyu Yao ^{1,2}

¹ College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China; 20202109042@stu.shzu.edu.cn (C.Y.); yaosiyu@stu.shzu.edu.cn (S.Y.)

² Key Laboratory Agriculture of Xinjiang Production and Construction Groups, Shihezi 832003, China

³ School of Information and Communication Engineering, Hainan University, Haikou 570228, China; zengtiwei@hainanu.edu.cn

⁴ Mechanical and Electrical Engineering College, Hainan University, Haikou 570228, China; 994330@hainanu.edu.cn (H.Z.); 994026@hainanu.edu.cn (W.F.);

* Correspondence: wl_mac@shzu.edu.cn

Abstract: Maize small leaf spot (*Bipolaris maydis*) is one of the most important diseases of maize. The severity of the disease cannot be accurately identified, the cost of pesticide application increases every year, and the agricultural ecological environment is polluted. Therefore, in order to solve this problem, this study proposes a novel deep learning network DISE-Net. We designed a dilated-inception module instead of the traditional inception module for strengthening the performance of multi-scale feature extraction, then embedded the attention module to learn the importance of interchannel relationships for input features. In addition, a dense connection strategy is used in model building to strengthen channel feature propagation. In this paper, we constructed a data set of maize small leaf spot, including 1268 images of four disease grades and healthy leaves. Comparative experiments show that DISE-Net with a test accuracy of 97.12% outperforms the classical VGG16 (91.11%), ResNet50 (89.77%), InceptionV3 (90.97%), MobileNetv1 (92.51%), MobileNetv2 (92.17%) and DenseNet121 (94.25%). In addition, Grad-Cam network visualization also shows that DISE-Net is able to pay more attention to the key areas in making the decision. The results showed that the DISE-Net was suitable for the classification of maize small leaf spot in the field.

Keywords: small spot disease; disease grading; dilated inception; multi-scale feature extraction; deep learning



Citation: Yin, C.; Zeng, T.; Zhang, H.; Fu, W.; Wang, L.; Yao, S. Maize Small Leaf Spot Classification Based on Improved Deep Convolutional Neural Networks with a Multi-Scale Attention Mechanism. *Agronomy* **2022**, *12*, 906. <https://doi.org/10.3390/agronomy12040906>

Academic Editors: Saeid Homayouni, Yacine Bouroubi and Karem Chokmani

Received: 9 March 2022

Accepted: 8 April 2022

Published: 9 April 2022

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



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

1. Introduction

Maize (*Zea mays* L.) is one of the four major food crops in China, and it is also an important feed crop and industrial raw material. Maize plays an important role in the development of the agricultural economy in China [1]. Hainan province has typical tropical climate resources, year-round climate and soil conditions suitable for seed breeding and production. Yazhou District, Sanya City, Hainan Province is a breeding base in the south of China, where maize can be sown all year round. In the north, due to climate and other factors, maize cannot be normally produced in winter, but in order to continue breeding research, Hainan province has become China's natural breeding base. With the development of the maize industry, diseases and insect pests have had a serious impact on maize yield, posing a serious threat to food security. In the process of maize breeding, the early diagnosis and accurate detection of maize diseases and pests and the judgment of their incidence provide a basis for pest control. Maize small leaf spot disease, also known as maize spot disease or southern leaf blight of maize, mainly caused by *Bipolaris maydis*, is one of the serious diseases in maize production, reducing maize production by 15~20%, and more than 50% in serious cases [2]. The scope of the damage caused by maize small

leaf spot disease has been further expanded recently, and the degree of damage has also become heavier. All these constraints have made maize small leaf spot disease control more difficult. Therefore, accurate identification of disease types and disease grades of maize is the key to selecting disease resistant varieties and precise application of drugs, which has become an important goal of intelligent field production control. On the premise of not affecting the effect of disease prevention and control, accurately identifying diseases and spraying pesticides can save nearly 60% of the dosage [3]. The current disease identification method is based on the manual investigation of crop experts [4], which is labor intensive, inefficient and error prone.

In recent years, with the emergence of computer vision and deep learning, using image information technology to detect and quantify crop diseases has become a research hotspot [5]. Several traditional machine learning methods that have been successfully applied in the research of plant disease recognition mainly include: Naive Bayes Classifier [6], support vector machine (SVM) [7], K-means clustering algorithm [8] and artificial neural network (ANN). Aravind et al. applied a support vector machine classifier (SVM) to classify three maize diseases, common rust, *Cercospora* leaf spot, leaf blight and healthy leaves, and achieved the best average accuracy of 83.7% [9]. Hossain et al. proposed a plant leaf disease detection and classification technology based on the K-nearest neighbor (KNN) classifier to classify anthracnose, bacterial blight, leaf spot and canker of different plants. This method can successfully detect and identify the selected diseases with an accuracy of 96.76% [10]. Kahar et al. used artificial neural network (ANN) to detect three types of rice diseases, Bacterial Leaf Blight (BLB), Leaf Blast Disease (LBD) and Bacterial Sheath Blight (BSB). The accuracy result for the recognition is 74.21% [11]. The above traditional machine methods need to extract classification features manually, the process is complex and the accuracy is low.

Deep learning has made a breakthrough in plant disease recognition [12], and has better performance than traditional machine learning methods in the field of plant disease recognition and classification. Nie et al. proposed a strawberry verticillium wilt detection network, which applied an attention mechanism to feature extraction of a disease detection network, and had better performance than the traditional disease detection network. The detection accuracy of strawberry verticillium wilt was 99.95% [13]. Rangarajan used deep learning model VGG16 to classify five diseases of eggplant. The classification accuracy was 94.3%. Finally, the optimized model is deployed on smart phones, and the classification accuracy is 91.3% under experimental conditions [14]. Waheed et al. proposed an optimized dense convolution neural network structure for maize leaf disease recognition and classification, and the recognition accuracy of three maize leaf diseases reached 98.06% [15]. Ramcharan et al. applied transfer learning to train the deep convolution neural network to identify three diseases and two pests of cassava. The best model achieved an overall accuracy of 93% for data not used in the training process. [16]. Haider et al. used a decision tree and different deep learning models to identify and classify wheat diseases. Results of both algorithms were then verified by domain experts that improved the decision trees' accuracy by 28.5% and CNN accuracy by 4.3% (leading to 97.2%) [17]. Zeng et al. proposed a group multi-scale attention network for rubber leaf disease image recognition and constructed a rubber leaf disease data set, with an accuracy of 98.06% for rubber disease. The recognition accuracy of the model on PlantVillage data set is 99.43% [18].

The above research on plant disease identification and classification has achieved ideal results, but plant disease classification is usually difficult, and disease identification can only judge the type of disease and the choice of pesticides. On the one hand, excessive use of pesticides will have adverse effects on the environment and food safety. On the other hand, it will also affect the effect of disease prevention and control. Mi et al. proposed a new deep learning network C-DenseNet, which embeds the Convolution Block Attention Module (CBAM) into DenseNet. The network was used to classify the severity of wheat rust, and the test accuracy was 97.99%. The results show that C-DenseNet with an attention mechanism is suitable for field classification of wheat stripe rust [19]. Zhang et al. established a single

ear segmentation model based on a full convolution network (FCN), which can effectively realize the segmentation of small wheat ears in a field environment. The disease grade was calculated using the ratio of the disease spot to the whole wheat ear. The classification accuracy is 92.5%; it can effectively classify the head blight of Wheat Fusarium in the field environment [20]. Fang et al. proposed a leaf disease grade identification method based on the convolutional neural network (CNN). The focus loss function was used to replace the standard cross entropy loss function, and the Adam optimization method was used. Finally, leaf disease grade identification was performed on a database containing 10 types of disease leaf images for eight crops, and it yielded a recognition accuracy of 95.61% [21]. Wang et al. proposed a two-stage model integrating deeplabv3 + and u-net in cucumber leaf disease severity classification (dunet) under a complex background. The experimental results show that the model can gradually segment leaves and disease spots with a complex background so as to complete the classification of disease severity. The average classification accuracy of disease severity reached 92.85% [22].

While the network in the above research achieved relatively high accuracy in plant disease classification, the classification of maize small leaf spot infection is more difficult because the image representation of different infection degrees of maize small leaf spot is highly similar, and the background is complex and difficult to distinguish. This study proposes a DISE-Net for maize small leaf spot recognition and classification. The main contributions of this paper are summarized as follows:

1. By dividing the grade of maize small leaf spot into fine-grained image classification, this paper proposes a DISE-Net model to realize the classification and detection of maize small leaf spot.
2. A maize small leaf spot grade data set was collected. The data set contained 1268 maize leaves, which were divided into five categories, including infection grade 4 and healthy maize leaves. In the complex environment, the field maize disease data set is collected through various mobile devices, and then manually calibrated according to the disease classification standard.
3. Based on the above data sets, the proposed DISE-Net model conducted extensive comparative experiments with VGG16, ResNet50, InceptionV3, MobileNetv1, MobileNetv2 and DenseNet121 classic networks (for example, performance, attention visualization). The results show that the proposed DISE-Net is better than the traditional six networks.

2. Materials and Methods

This section describes in detail the materials and methods used in this study, including the collected image data set and the proposed DISE-Net model.

2.1. Overview of Sampling Area

The image samples required for the experiment were collected from the maize breeding experimental field (18°22' N, 109°11' E) of Nanbin farm, Sanya City, Hainan Province. The maize was planted in rows from north to south, with good light transmission, plant spacing of 2500 ± 10 mm and row spacing of 580 ± 10 mm. As shown in Figure 1, the maize variety ND414 in the experimental field was affected by rainfall in the typhoon period (typhoon period was from June to November) during the whole growth period. The rainfall is more than 500 mm, the temperature is between 23 and 30 °C, and the average air humidity is 80%, which is suitable for the breeding and development of maize small leaf spot. *Bipolaris maydis* is mainly harmful to maize leaves. The main manifestation of the disease is oval or fusiform spot on the leaves, with yellow brown in the center and purple or dark brown in the edge. When the air humidity is high, the surface of the spot is gray brown with sparse mildew layer, and then the leaves turn yellow and die. The disease occurrence degree of the experimental plot is suitable for the grading test requirements of disease grade.



Figure 1. Maize experimental field.

2.2. Image Acquisition

The data collection time is from 8:00 am to 5:00 pm. Different mobile devices (Nikon D90, iPhone 11, Vivo X60, Redmi K40) are used to obtain RGB images of maize leaves respectively to ensure the diversity of samples. Their pixel resolution is 4288×2848 , 4032×3024 , 4160×3120 , 3000×3000 . The camera or mobile phone camera is 12~60 cm away from the maize leaves, and the maize leaves with different disease grades are photographed at the upper, middle and lower positions of the same plant in a random way. The collected images have different weather and background, such as sunny, cloudy or cloudy days, with weeds or photographer's hands on the background. A total of 1268 images of maize leaves were collected. These images further divided the infection degree of maize small leaf spot into five grades through the inspection of maize pathologists. The division standard is shown in Table 1, and the grade image sample is shown in Figure 2.

Table 1. Disease classification standard.

Disease Grade	Symptoms
0	No obvious symptoms
1	The leaves appear chlorosis and scattered oval lesions ($0 \leq K < 0.1$)
2	The leaf spots became larger and some of them were spindle shaped and densely distributed ($0.1 \leq K < 0.3$)
3	The adjacent leaf spots overlap and wither at the edge ($0.3 \leq K < 0.5$)
4	More than 1/2 of the leaves showed wilting symptoms ($K \leq 0.5$)

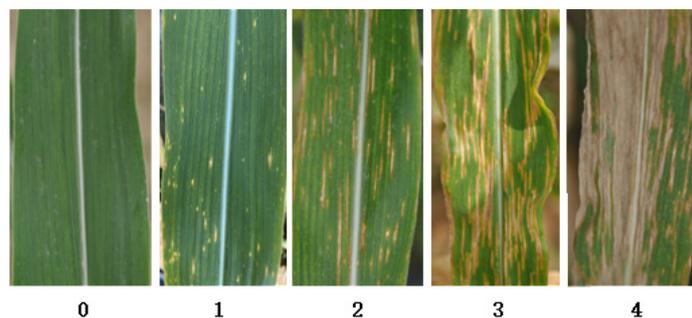


Figure 2. Level image sample.

K in Table 1 is calculated by experts on the spot based on experience, and K is the estimated value of the disease range. The calculation formula is:

$$K = \frac{A_1}{A} = \frac{N_1}{N} \quad (1)$$

where, A is the whole leaf area of the crop, A_1 is the lesion area, N is the number of pixels of the whole leaf, and N_1 is the number of pixels of the lesion area [21]. The disease grade is determined by ratio K .

Most of the pictures of maize small leaf spot were taken in the tasseling stage of maize, and some maize plants were seriously affected by disease, with many leaves of grade 4 disease. Table 1 shows the ratio of leaf spot area to total leaf area of maize. Grade 0: Healthy leaves without obvious symptoms; Grade 1: Leaf Chlorosis and scattered oval lesions, accounting for less than 10% of the leaf area; Grade 2: The disease spots of the leaves become larger and distributed densely in a spindle shape, accounting for 10~30% of the leaf area; Grade 3: The adjacent disease spots of the leaves overlap and the edges of the leaves wither, accounting for 30~50% of the leaf area; Grade 4: There are a large number of disease spots on the leaves, accounting for more than 50% of the leaf area, and there are large areas of withering and death.

2.3. Image Enhancement

The ROI region of the original RGB image is clipped, and the image preprocessing operations such as image filtering and denoising are carried out. The size of the sample image is uniformly adjusted to 224×224 pixels; this is used as the input of image analysis to reduce the calculation time and improve the efficiency of image processing. The constructed original data set includes 112 healthy samples of maize small leaf spot, 125 first-class diseases, 328 second-class diseases, 348 third-class diseases and 355 fourth-class diseases.

The number of samples in the constructed data set is small, and there is an imbalance in the number of images with different levels of diseases. In deep learning, the small number of samples and uneven distribution will affect the accuracy of model recognition [23]. Data enhancement technology can increase the diversity of samples [24]. The data enhancement strategies in this paper include color enhancement, random rotation, random clipping and horizontal random flip. The example of sample data enhancement is shown in Figure 3. Including brightness adjustment, contrast adjustment, saturation adjustment and chroma adjustment. Random rotation refers to the random rotation of the picture. The rotation angle range is -30° ~ 30° . Random clipping refers to the arbitrary clipping of a part of the picture between 0.6 and 1.0. Horizontal random flipping refers to the random flipping of the picture into a mirror image which is then converted into 224×224 picture size. The data enhancement strategies in this paper were randomly performed with a probability of 50%. The enhanced data set was as follows: 1172 healthy cases of maize small spot disease, 1139 cases of first-level disease, 1689 cases of second-level disease, 1540 cases of third-level disease and 1987 cases of fourth-level disease. The detailed results of the data set before and after the application of the enhanced treatment are shown in Table 2.

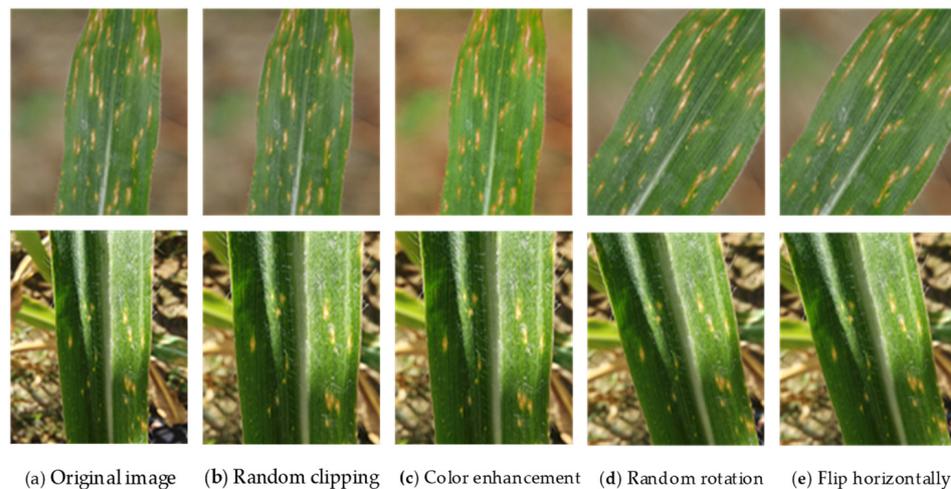


Figure 3. Example of data enhancement.

Table 2. Data sets before and after image enhancement processing.

Disease Grade	Images	Images (Augmentation)	Training Images	Validation Images	Testing Images
0	112	1172	702	235	235
1	125	1139	689	225	225
2	328	1689	1013	338	338
3	348	1540	924	308	308
4	355	1987	1027	480	480

2.4. Architecture of DIS-NET Model

2.4.1. Network Architecture

The design basis of this model is inspired by GoogLeNet. In order to enhance the extraction ability of fine-grained disease features of maize, the multi-scale dilated-inception module is designed combining the inception module and dilated convolution. Finally, the dense connection strategy is adopted to strengthen the connection between features. The hierarchical detection network structure of maize small leaf spot is shown in Figure 4. The DISE-Net model includes three parts: The first part is the “Pre-Network Module” and consists of 7×7 and 3×3 convolution layers with the number of convolution kernels being 96 and 192, respectively, followed by a 3×3 max-pooling layer and a batch normalization layer. The second part is the “Cascaded Dense DISE Module”, composed of three DISE modules with dense connections. The DISE structure consists of a dilated-inception block and a squeeze-and-excitation networks block. The last part is composed of a convolution layer, an average pooling layer, a fully connected layer and a five-way Softmax layer. Moreover, the batch normalization layer and the ReLu activation function are added after each convolution layer to accelerate the convergence of the model. The main parameters of the DISE-Net model are shown in Table 3.

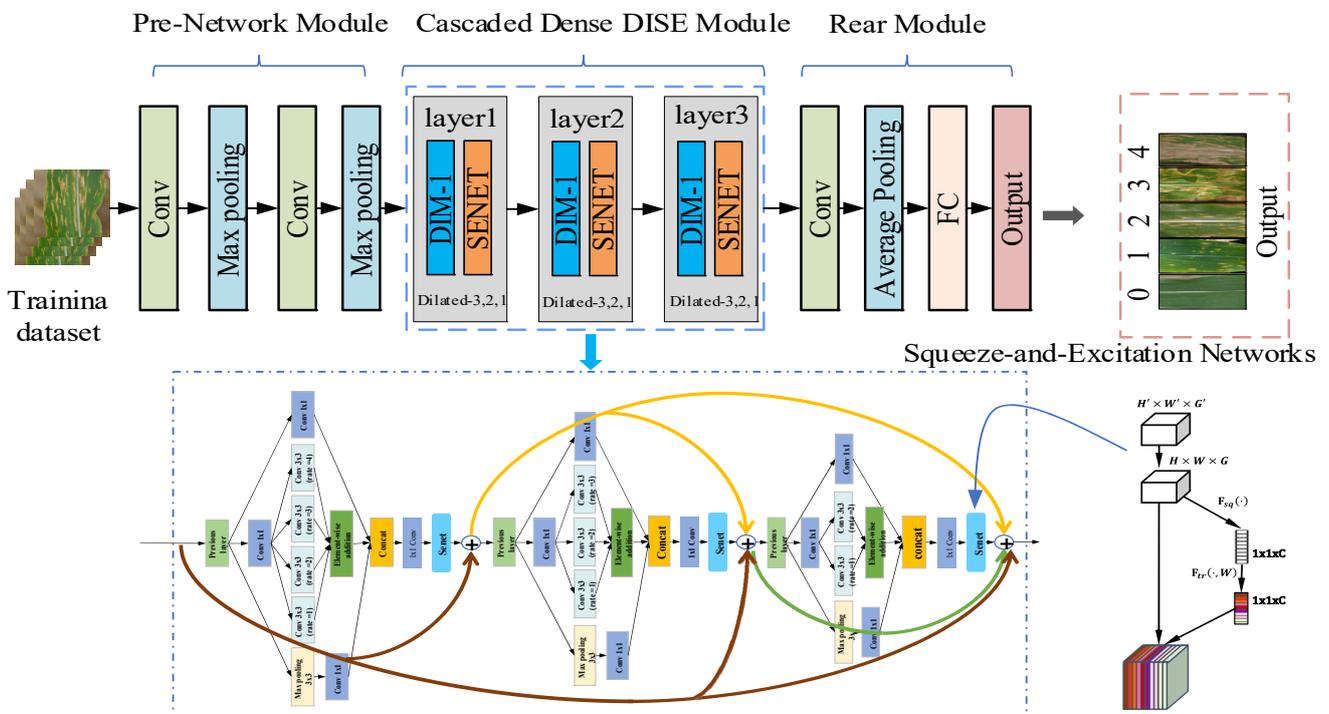


Figure 4. Neural network structure for grading detection of maize small leaf spot.

Table 3. Main parameters of DISE-Net model.

Layer	Patch Size/Stride	Output Tensor
Input	Augmented images	$224 \times 224 \times 3$
Convolution	$7 \times 7/2$	$112 \times 112 \times 96$
Max_pooling	$3 \times 3/2$	$56 \times 56 \times 96$
Convolution	$3 \times 3/1$	$56 \times 56 \times 192$
DISE Block1	$3 \times 3/\text{dilated } 3, 2, 1$	$56 \times 56 \times 288$
Max_pooling	$2 \times 2/2$	$28 \times 28 \times 288$
DISE Block2	$3 \times 3/\text{dilated } 3, 2, 1$	$28 \times 28 \times 360$
DISE Block3	$3 \times 3/\text{dilated } 3, 2, 1$	$28 \times 28 \times 360$
Max_pooling	$2 \times 2/2$	$14 \times 14 \times 360$
Convolution	$3 \times 3/2$	$7 \times 7 \times 256$
Average_Pooling	-	$1 \times 1 \times 256$
Softmax	-	5

2.4.2. Multi-Scale Dilated-Inception Module

Maize small leaf spot infections grade is more difficult because the difference is usually small under different levels of maize small leaf spot infection, which makes it difficult for the traditional model to identify the images with subtle differences.

As shown in Figure 3a,b, the spots of grade 1 disease are small and scattered, and the color of healthy leaves is similar to that of grade 1 disease. As shown in Figure 3c,d, the disease spots of grade 2 disease and grade 3 disease become larger and in spindle shape. The difference lies in the density and overlap of disease spots. Fine grain characteristics (color and texture of disease) need to be considered in identifying these grades. As shown in Figure 3e, more than 1/2 of the leaves have wilting symptoms when the maize small leaf spot disease is grade 4, which is quite different from other grades. The coarse grain characteristics (lesion size and texture) need to be considered in identification. Therefore, the identification of maize small leaf spot should consider both coarse grain characteristics and fine grain characteristics.

The main idea of dilated convolution is to insert holes (zero) in the standard convolution kernel to increase the receptive field [25]. Dilated convolution is defined as:

$$(F *_l k)(p) = \sum_{s+lt=p} F(s)k(t) \quad (2)$$

where F is a discrete function and k is a discrete filter of size $(2r + 1)^2$, where $*_l$ is called a dilated convolution or an l -dilated convolution. As shown in the middle part of Figure 5, k is a 3×3 filter, and the kernel dilation rates are 1, 2, 3, respectively. Different dilation rates of the dilated convolution can extract different receptive field features, and multi-scale receptive field features can be extracted by fusing multiple dilatation convolution with different dilation rates [26]. Therefore, the combination of small expansion rate and large expansion rate can not only increase the width of the model, but also improve the performance of multi-scale feature extraction.

The inception block in GoogLeNet has been applied [27], as shown in Figure 6a. The inception block is a multi-branch structure composed of convolution layers with multiple convolution kernel sizes parallel. This structure can not only effectively extract multi-scale features, but also reduce model parameters.

The proposed dilated-inception module is inspired by the inception module and dilated convolution. The principal idea of the dilated-inception module is to utilize multiple dilated convolutional layers with different dilation rates, working as multi-scale feature extractors with various receptive field sizes, as shown in Figure 5. The small dilation rate extracts the features of the fine-grained lesion, and the large dilation rate pays more attention to the features of the coarse-grained lesion. The dilated-inception module can simultaneously extract the features of fine-grained lesions and coarse-grained lesions.

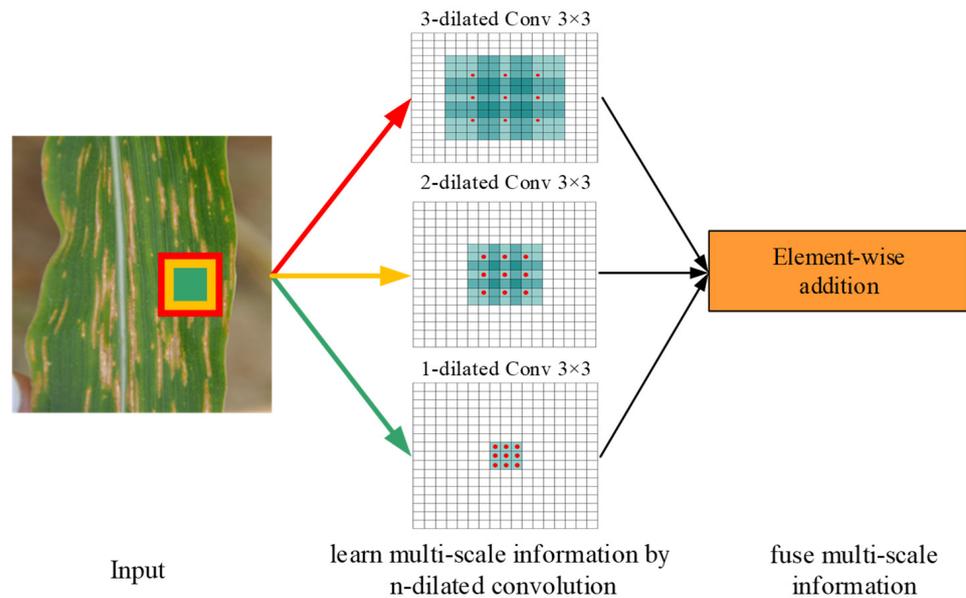


Figure 5. Multi-scale dilation convolution; how to learn multi-scale feature information.

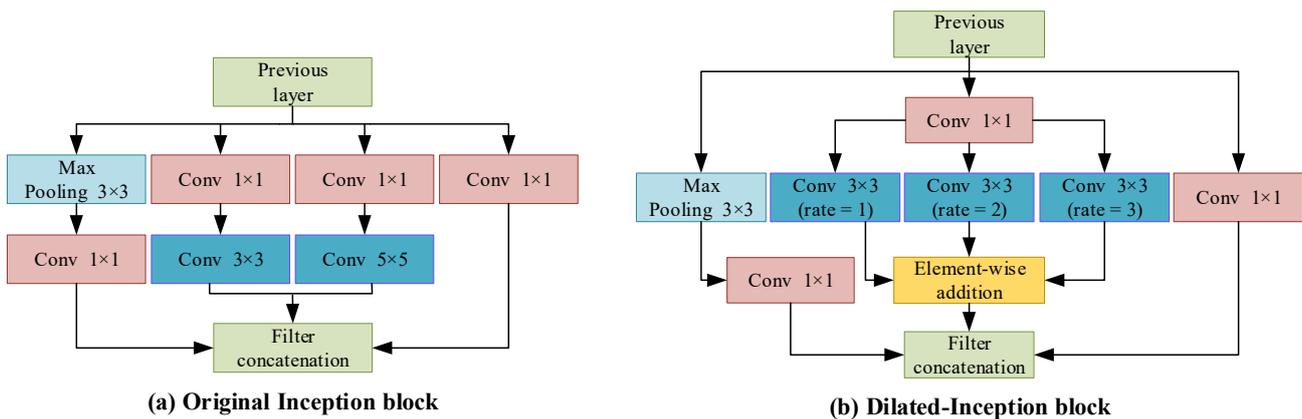


Figure 6. The inception module with its variations.

In this paper, we utilize this module to capture multi-scale spatial information. As shown in Figure 6b, our dilated-inception module can be viewed as a variant of the GoogleNet with a combination of three parallel dilated convolutions inside. The dilated-inception module consists of a 3×3 maxpooling with a 1×1 convolution, a 1×1 convolution, and three dilated convolution with kernel sizes of 3×3 and dilation rates of 1, 2 and 3. In addition, on the one hand, we use a single 1×1 top convolution for dimension reduction in the three parallel dilated convolution. On the other hand, the three dilated convolutions use element-wise addition for filter fusion for dimension reduction and multi-scale feature fusion.

2.4.3. Squeeze-and-Excitation Networks

Recently, attention mechanism has been widely used, including image processing [28,29], speech recognition [30], natural language processing [31], etc. The attention mechanism pays attention to the useful information of various channels of the network, inhibits the useless information, and effectively improves the performance of the existing state-of-the-art model with a slight computation cost. SE-Net introduced the SE module [32]. As shown in Figure 7, it dynamically selects features by learning the importance of different channel features, and SE-Net is flexible and can be directly applied to the existing network. The SE-Net module includes compression, excitation and reconstruction.

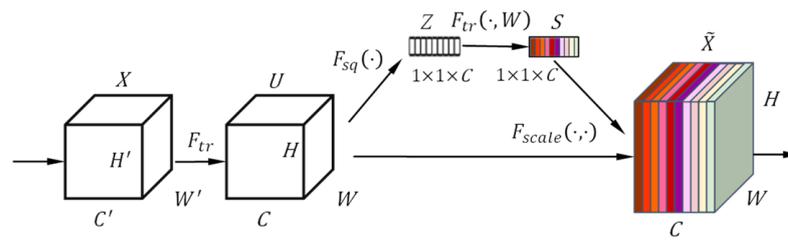


Figure 7. Squeeze-and-Excitation Networks (SE-Net).

In the compression step, namely F_{sq} , The feature map U with C channels is compressed into vector Z of $1 \times 1 \times C$ by the global average pool layer. The specific formula is as follows:

$$Z_c = F_{sq}(U_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H U_c(i, j) \tag{3}$$

Secondly, In the excitation step, namely F_{ex} , the global feature S is obtained through two fully connected layers, one ReLU activation layer and one Sigmoid layer, respectively. The specific formula can be described as:

$$S = F_{ex}(Z, W) = \sigma(g(Z, W)) = \sigma(W_2 \vartheta(W_1 Z)) \tag{4}$$

Vector S contains the information of the weight coefficients of different channel features. Where σ and ϑ are Sigmoid activation function and ReLU activation function respectively, $W_1 \in R^{\frac{C}{r} \times C}$ and $W_2 \in R^{C \times \frac{C}{r}}$ are dimension reduction and restoration parameters, respectively. r is the reduction factor, which is set to 16 in this paper.

Finally, in the reconstruction stage, namely F_{scale} , S_c with C weight coefficients and feature map U are weighted to obtain the weighted feature map \tilde{X} , each channel has different weight coefficients to express the importance of feature information.

$$\tilde{X}_c = F_{scale}(U_c, S_c) = S_c U_c \tag{5}$$

The DISE module consists of a Dilated-Inception block and SE-Net block, as shown in Figure 8.

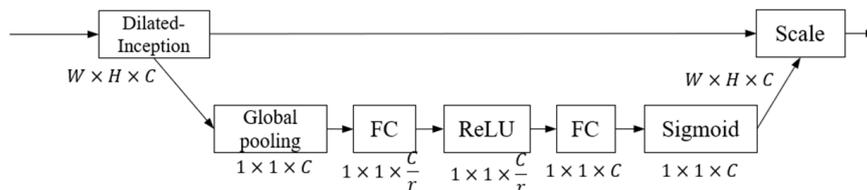


Figure 8. The DISE module consists of a dilated-inception block and an SE-Net block.

2.4.4. Cascade Dense Connectivity Strategy

With the deepening of the network layer, the vanishing-gradient problem becomes more serious, and it is difficult to transmit the feature of the low-level network to the deeper part of the network, which seriously affects the recognition performance of the network. The DenseNet network takes feature maps of all previous layers in the network serve as input to all subsequent layers [33], and all feature maps are connected by concatenating, which improves the connection between different feature layers and effectively prevents the vanishing-gradient problem. The specific formula of the dense connectivity strategy is as follows:

$$x = H_\lambda([x_0, x_1, \dots, x_{\lambda-1}]) \tag{6}$$

where $[x_0, x_1, \dots, x_{\lambda-1}]$ denotes the concatenation of the maps from the previous layers.

The dense connection strategy in this paper is to cascade three DISE modules to realize feature reuse, and then adopt dense connection to strengthen feature propagation among multiple feature layers and improve parameter utilization, as shown in Figure 9.

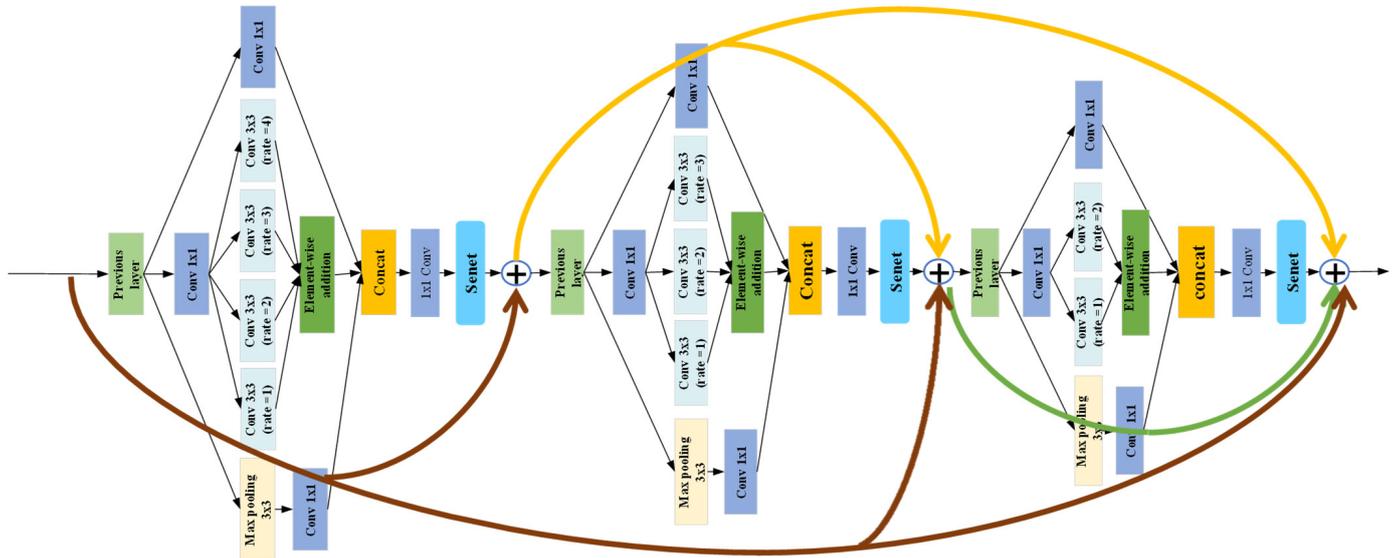


Figure 9. The overview of the cascade dense connectivity strategy.

2.5. Overall Process

The overall process mainly includes three stages. Firstly, the sample images of maize leaves were obtained from the maize experimental field. According to the knowledge of domain experts, the image of maize leaf disease is graded and labeled. Secondly, the original image is processed by image preprocessing, image clipping, image filtering and image denoising. Next, the image data enhancement scheme is used to synthesize the new image, increase the image and balance the category of the data set to avoid the over fitting problem. Including color enhancement, random rotation, random clipping and horizontal random flip. Finally, the enhanced sample image is used as the training input to train the proposed DISE-Net. The overall process is shown in Figure 10.

2.6. Evaluation Indexes

In this study, we analyze the output results of different models in detail. As defined in Equations (7)–(10), precision, recall, $F1_score$ and accuracy are used as evaluation indexes to comprehensively evaluate the performance of the deep learning algorithm:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1_score = \frac{2TP}{2TP + FP + FN} \quad (9)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (10)$$

where TP , TN , FP and FN are the number of true-positive samples, true-negative samples, false-positive samples and false-negative samples, respectively. Precision estimates how many of the predicted positive samples are positive. Recall is the assessment of how many of all positive samples can be correctly predicted as positive. Accuracy is the most intuitive indicator to measure the quality of a model. Model size and parameters are often used to measure model complexity.

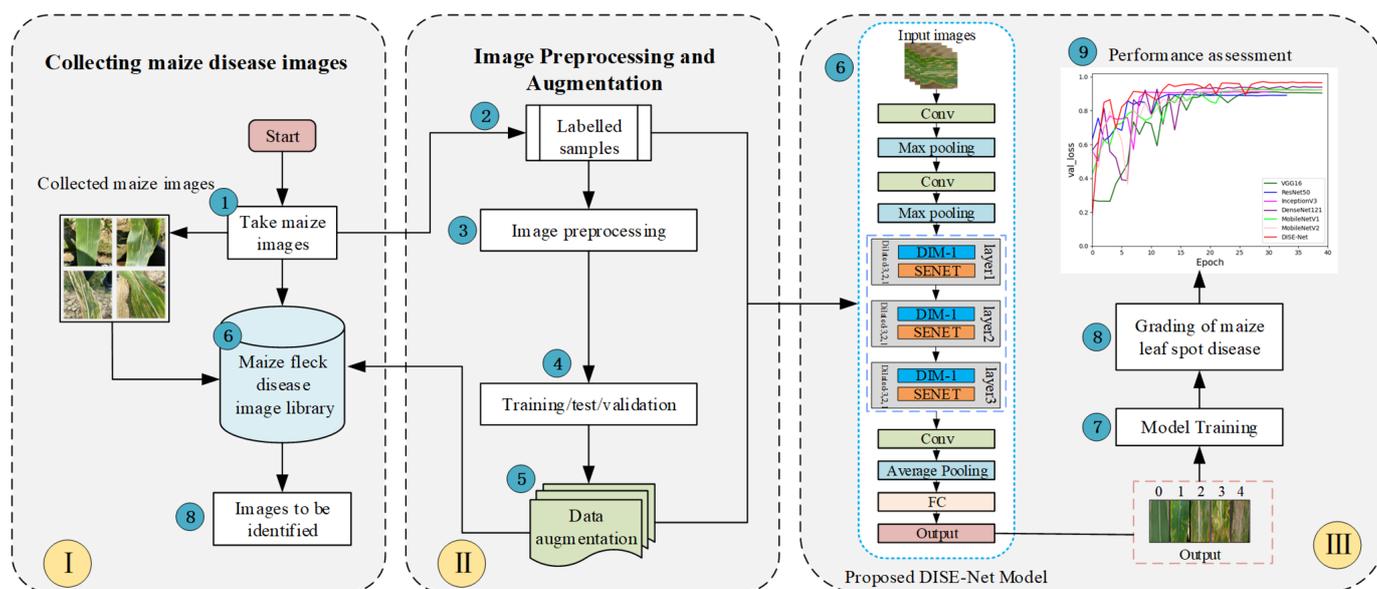


Figure 10. Whole process of grading detection of maize small leaf spot disease.

3. Results and Discussion

3.1. Experimental Setup

Data augmentation and deep learning algorithms are implemented in Keras deep learning framework based on CNN using python language. The experimental hardware configurations include an Intel i5-10400F CPU (2.90 GHz), a memory of 16 GB and an RTX 2060S graphics card.

Considering the performance of hardware equipment, the enhanced maize small leaf spot data set is divided into three groups: Training set, verification set and test set; these make up 60%, 20% and 20% respectively. The batch size and the number of iterations for all network models are 16 and 40, respectively, and categorical crossentropy is used to optimize the model. Stochastic Gradient Descent (SGD) was adopted for training. The initial learning rate is set to 0.1 for the first epoch, and the learning rate is dynamically adjusted by using Keras’s ReduceLROnPlateau function. If the accuracy of the validation set does not improve after three iterations, the learning rate will be reduced by half. In this way, training time can be shortened compared to fixed learning rate. The epoch parameter was selected based on training experience. These parameters are used to test all deep learning models. The specifications of training parameters of optimization methods are presented in Table 4.

Table 4. Specification of training parameters.

Parameter	Value
Optimizer	Stochastic Gradient Descent
Loss function	Categorical crossentropy
Batch size	16
Epoch	40
Initial learning rate	0.1
Momentum	0.9
Patience (ReduceLROnPlateau)	3
Factor (ReduceLROnPlateau)	0.5

3.2. Performance Comparison of Different Models

In order to verify the effectiveness of the neural network model, based on the disease data set we constructed, and according to the training parameters in Section 2.5, the DISE-Net model is compared with VGG16, ResNet50, InceptionV3, MobileNetV1, MobileNetV2

and DenseNet121 classic networks. Figure 11 shows the accuracy curve and loss curve of the above six networks and DISE-Net on the training and validation data set. It can be seen from the accuracy curve that the DISE-Net model has the highest recognition accuracy and the fastest convergence speed on the maize small leaf spot data set, which verifies the effectiveness of the improved model. However, the loss curve shows that the loss value of the model is relatively general.

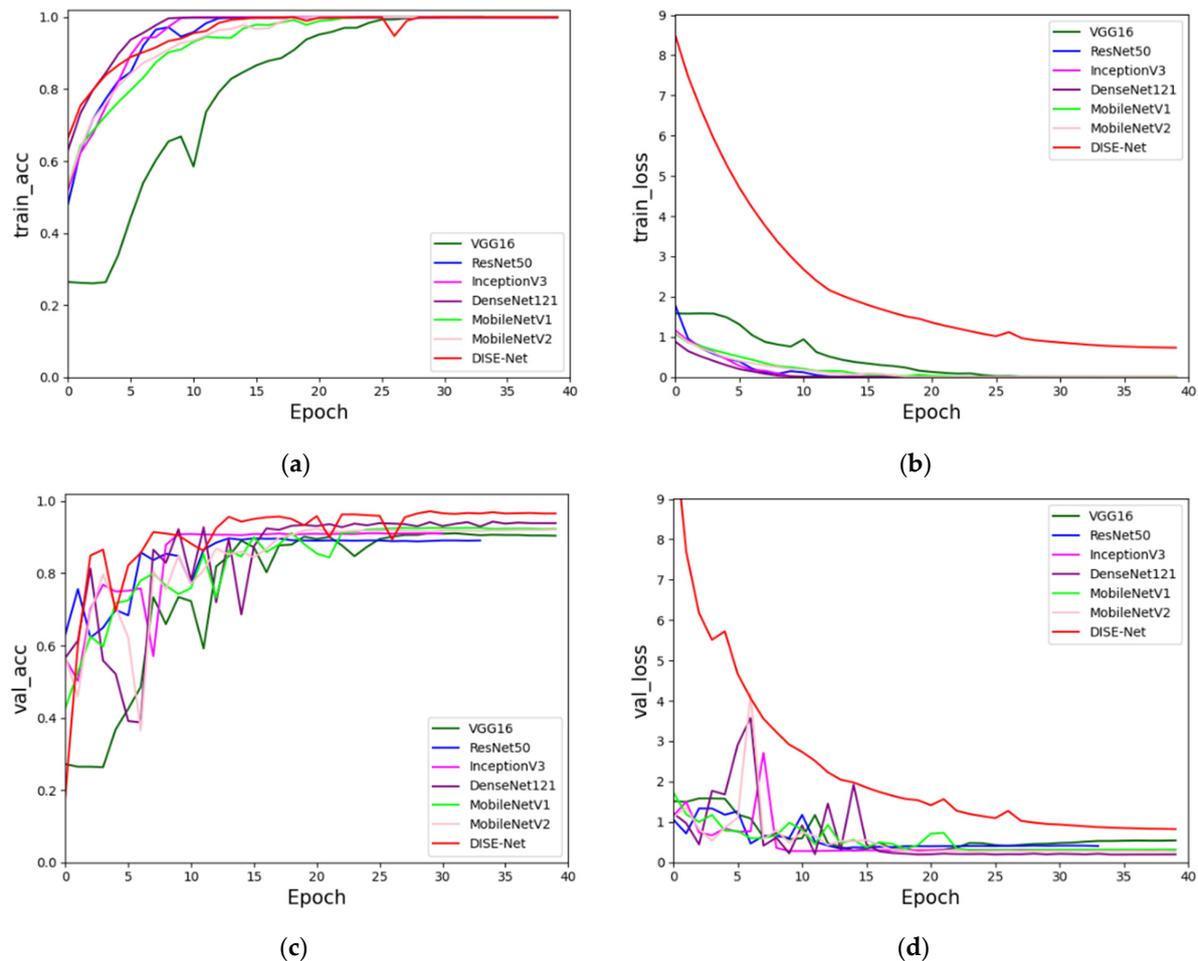


Figure 11. Accuracy curve and loss curve (a) Training accuracy curve; (b) Training loss curve; (c) Validation accuracy curve; (d) Validation loss curve.

As shown in Table 5, the size, parameters, precision, recall rate, $F1_score$ and accuracy of the seven networks are compared. The results show that the DISE-Net model is significantly better than the other six models in all evaluation indexes. The recognition accuracy of InceptionV3 and ResNet50 is relatively low, 90.97% and 89.77%, respectively. Compared with the previous two models, the recognition accuracy of DenseNet121 is higher, reaching 94.25%, and the recognition accuracy of the proposed model reaches 97.12%. On the premise of a slight increase in the accuracy of the proposed model compared with DenseNet121, the size of the model is reduced 11.9 times, 12.9 times and 3.9 times, respectively, compared with the network of InceptionV3, ResNet50 and DenseNet121. Compared with mobilenetv1 and mobilenetv2 lightweight networks, the accuracy of the model is improved by 4.61 and 4.95 percentage points, respectively, which is the significant advantage of the model in multi-scale feature extraction and feature reuse. In a word, among the seven compared models, the DISE-Net model has the smallest parameters, which can obtain better convergence and the highest accuracy.

Table 5. Test results of grading detection of maize small leaf spot.

Models	Size (MB)	Parameters (M)	Precision (%)	Recall (%)	F1_Score (%)	Accuracy (%)
VGG16	1000	134	88.01	87.94	87.99	91.11
ResNet50	180	23.6	90.37	90.33	90.60	89.77
InceptionV3	167	21.8	89.88	89.91	89.97	90.97
MobileNetv1	24.8	3.23	92.56	92.92	92.78	92.51
MobileNetv2	17.8	2.28	92.28	92.41	92.11	92.17
DenseNet121	54.6	54.6	95.34	95.48	95.34	94.25
DISE-Net	14	1.69	96.25	96.28	96.08	97.12

In addition, DISE-Net also uses a confusion matrix, as shown in Figure 12. It can be seen from the confusion matrix that misclassification mainly occurs between levels 2 and 3, and between levels 3 and 4, which is mainly caused by the very small difference of disease spots between them, the chaotic field background and inconsistent light intensity. Therefore, some misjudgments are tolerable. The diagnostic accuracy of healthy leaves (0_HL), first-class diseases (1_BL) and second-class diseases (2_BL) of maize was more than 95.6%, and the diagnostic accuracies of third-class diseases (3_BL) and fourth-class diseases (4_BL) of maize small spot disease were 94.5% and 95.2%, respectively. Table 6 lists the performance of identifying different disease levels in terms of accuracy, precision, recall and F1_score. The results show that the average values of accuracy, precision, recall and F1_score are 96%. This result proves the effectiveness of the model in the classification of maize small leaf spot.

Table 6. Classification report.

	Precision (%)	Recall (%)	F1_Score (%)	Support
0_HL	0.99	1.00	0.99	235
1_BH	0.98	0.96	0.97	228
2_BH	0.96	0.96	0.96	337
3_BH	0.92	0.94	0.93	307
4_BH	0.97	0.95	0.96	397
Macro avg	0.96	0.96	0.96	1504
Weighted avg	0.96	0.96	0.96	1504
Accuracy	-	-	0.96	1504

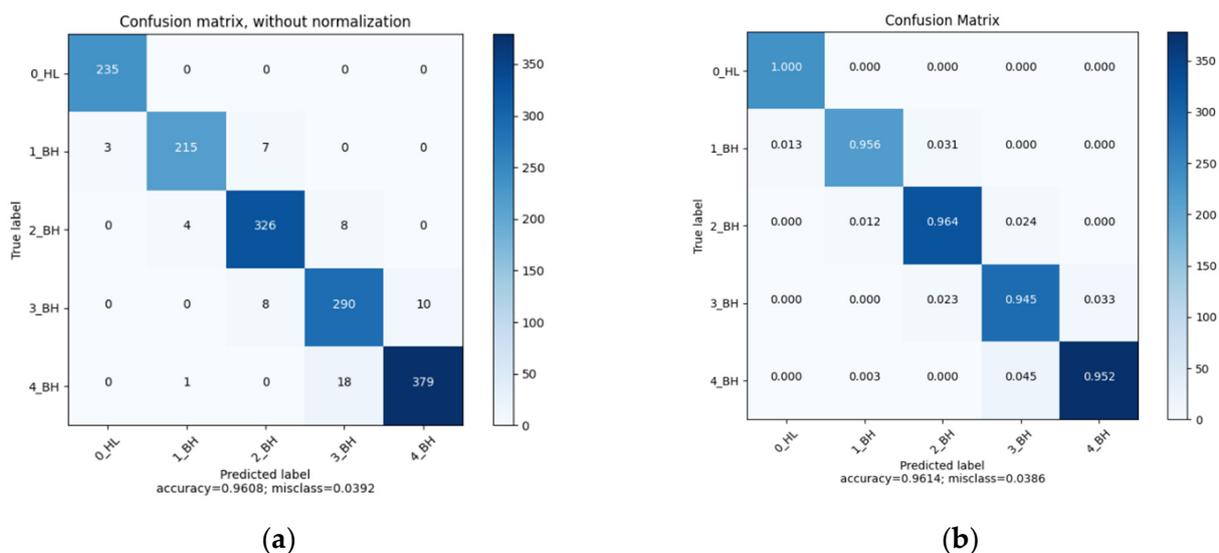


Figure 12. Confusion matrix of DISE-Net. (a) Confusion matrix without normalization. (b) Normalized confusion matrix.

3.3. Ablation Experiments

In order to explore the influence of the dilation rate of the dilated-inception module on the model, the following four combinations were designed: DISE_V1(D = 1), DISE_V2(D = 2), DISE_V3(D = 3) and DISE_V4(D = 4); D represents the dilation rate of the dilated convolution in the cascaded dense DISE module. The results are shown in Figure 13. In the beginning, the accuracy increases with the increase of the dilation rate, and the DISE_V3 has a better effect when the dilation rate is set to 3; the accuracy is 97.98%. However, when the dilation rate increases to 4, the accuracy of DISE_V4 reaches 97.03, which is 0.33 percentage points lower than that of DISE_V3.

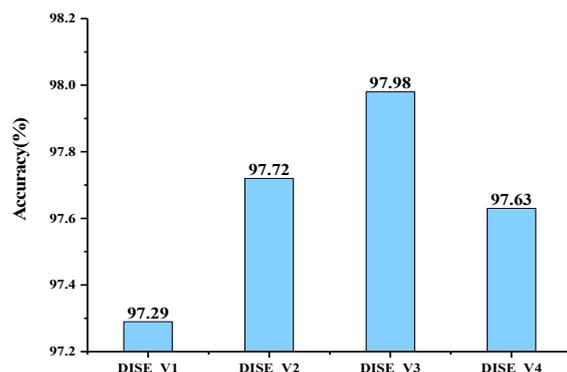


Figure 13. The effect of different dilated rates on the accuracy of the convolution neural network.

It can be seen from Table 7 that when the dilation rate corresponding to the dilated inception module is set to 3, although the model size is 5.6 M larger than that when the dilation rate is set to 1, the accuracy is improved by 0.69%. It should be noted that the dilation rate of dilated convolution should be properly selected. Different receptive field features can be extracted with different dilation rates. Excessive dilation rate of dilated convolution may cause parameter redundancy, computational resource waste and precision decline due to overfitting problems. If the dilation rate of dilated convolution is too small, the classification result will be unsatisfactory due to insufficient coarse-grained feature extraction. The appropriate inflation rate not only improves the recognition rate, but also reduces the model size.

Table 7. The operation effect of network structure under different dilation rate combinations.

Models	Network Architecture	Size (MB)	Parameters (M)	FLOPs (M)	Accuracy (%)
DISE_V1	Dilation rate = 1	17.5	2.27	5.08	97.29
DISE_V2	Dilation rate = 2	20.3	2.63	6.01	97.72
DISE_V3	Dilation rate = 3	23.1	2.98	6.93	97.98
DISE_V4	Dilation rate = 4	25.9	3.34	7.86	97.63

In order to further improve the recognition performance of the model, the SE-Net attention module is added after each dilated-inception module, which is combined into the DISE-Net module. As shown in Figure 14, Base represents the original DISE-Net model, and Base(no_SE-Net) represents the model after removing the SE-Net module. Table 8 shows the specific impact of different modules on the model. Obviously, after adding SE-Net, the model size is increased by 2.7 M, the model recognition accuracy is improved from 94.91% to 97.12% and the performance in accuracy, precision, recall and *F1_score* are significantly improved. The results show that the SE-Net is helpful for learning the useful information of each channel of the network and improving the recognition accuracy of the model.

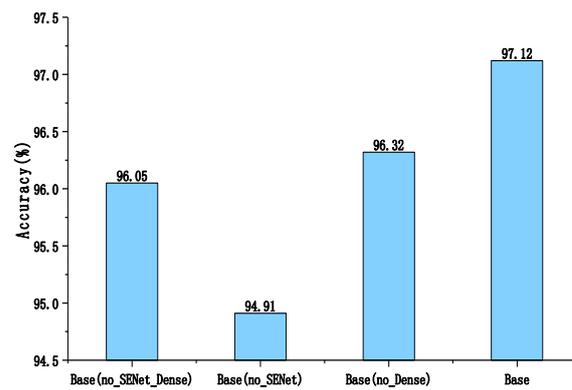


Figure 14. The effect of SENet and dense connection strategy on the accuracy of the convolution neural network.

Table 8. The effect of different modules on the model.

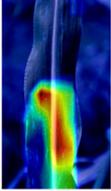
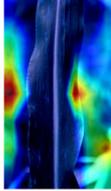
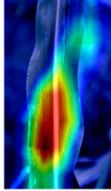
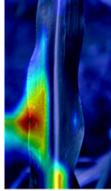
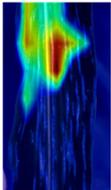
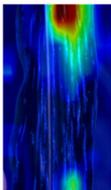
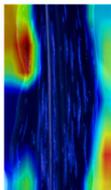
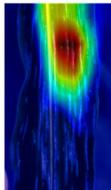
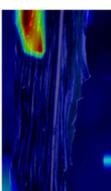
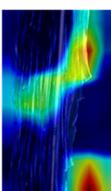
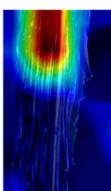
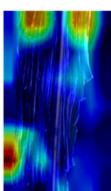
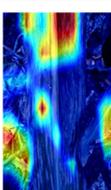
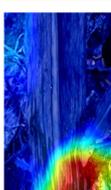
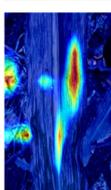
Models	Size (MB)	Parameters (M)	Precision (%)	Recall (%)	F1_Score (%)	Accuracy (%)
Base(no_SENet_Dense)	11.3	1.44	95.13	95.54	95.22	96.05
Base(no_SENet)	13.9	1.79	95.42	95.67	95.42	94.91
Base(no_Dense)	11.3	1.44	95.67	95.88	95.55	96.32
Base	14	1.69	96.25	96.28	96.08	97.12

This experiment evaluates the effect of the dense connection strategy on the model. As shown in Figure 14, Base(no_Dense) represents the model after removing the dense connection strategy, and Base(no_SE-Net_Dense) represents the model after removing the SE-Net module and dense connection strategy. As shown in Table 8, the results show that the model recognition rate after removing the dense connection strategy is 96.32%, and the model recognition rate after removing the senet module and dense connection strategy is 96.05%. Compared with the original DISE-Net model, the recognition accuracy decreased by 1.07% and 0.80%, respectively. The performance in accuracy, precision, recall and *F1_score* is also inferior to the original DISE-Net model. The above experimental results show that the dense connection strategy can enhance the feature propagation between feature layers with different depths and improve the recognition accuracy of the model.

3.4. Network Attention Visualization

In order to better understand the learning capacity of the proposed DISE-Net model, Grad-CAM is used to visualize the classification results. In particular, the randomly selected samples were tested. As shown in Table 9, the attention heat maps of four models for identifying different grades of maize leaf spot disease are shown. The visualization results were superimposed by the image of maize small spot disease and the heat map of maize disease. The heat maps of DenseNet121 and ResNet50 both highlight localized areas of small plaque, but the accuracy of the heat maps is not high. The MobileNetv2 heat map highlights some of the patch areas, but contains a lot of extraneous background information. Compared with DenseNet121, MobileNetv2 and ResNet50 models, the propose DISE-Net model can accurately focus on the key areas of maize small spot disease, has high accuracy in heat map and has minimal attention to irrelevant complex backgrounds, thus achieving higher disease identification accuracy than other models.

Table 9. Attention heat maps of DenseNet121, MobileNetv2, ResNet50 and DISE-Net for identifying maize minor spot disease with different grades.

Disease Grade	Input Image	DenseNet121	MobileNetv2	ResNet50	DISE-Net
1					
2					
3					
4					

4. Conclusions and Future Work

In variety breeding, disease resistance to growth and development in the process of disease severity classification assessment is of great significance. The differences between different grades of maize leaves are usually subtle, which brings great challenges to image classification. Therefore, in this study, the maize small spot disease grading data set was constructed, and we proposed a DISE-Net structure for the maize small spot disease-grading task. In DISE-Net, the dilated-inception module is designed to aggregate multi-scale context information and improve the identification accuracy of the model for different disease grades. At the same time, it uses the SE-Net attention module after each dilated-inception module in the model to learn the importance of interchannel relationships for input features. In addition, the model also adopts a dense connection strategy to strengthen the propagation of channel characteristics and improve the performance of the infection grade scoring task. The recognition accuracy of this model reaches 97.12%, which is the best in the comparative model. The results of this study realized the end-to-end classification of maize small spot disease, and provided a solution and reference for the application of the deep learning method in crop disease classification. There are still deficiencies in this study: (1) These data are obtained in the experimental field of maize small spot disease breeding. Therefore, there are few pictures of other diseases or multiple diseases on a leaf. (2) This study does not consider the identification of multiple maize diseases. At present, the network is only suitable for the classification and identification of maize small spot disease grade. In future work, we plan to apply the model to portable devices to widely monitor and identify maize disease information. At the same time, we will apply the model to other disease classifications.

Author Contributions: Conceptualization, C.Y. and T.Z.; Data curation, C.Y. and H.Z.; Formal analysis, C.Y. and W.F.; Funding acquisition, W.F. and L.W.; Investigation, C.Y., T.Z. and H.Z.; Methodology, C.Y. and S.Y.; Project administration, L.W. and W.F.; Resources, C.Y. and T.Z.; Software, C.Y. and H.Z.; Supervision, W.F.; Validation, C.Y., W.F., T.Z., L.W. and S.Y.; Visualization, C.Y. and T.Z.; Writing—original draft, C.Y. and L.W.; Writing—review and editing, C.Y. and T.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Key R&D projects in Hainan Province (Grant No. ZDYF2020042).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on-demand from the first author at (wl_mac@shzu.edu.cn).

Acknowledgments: The authors would like to thank their schools and colleges, as well as the funding of the project. All supports and assistance are sincerely appreciated. Additionally, we sincerely appreciate the work of the editor and the reviewers of the present paper.

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

References

1. Wang, Q.; Chen, Y. Advantages Analysis of Corn Planting in China. *J. Agric. Sci. Tech.* **2018**, *20*, 9.
2. Gao, G. Discussion on Hazard Symptoms and Prevention Methods of Corn Southern Leaf Blight. *J. Agric. Catas* **2016**, *6*, 3.
3. Jensen, H.; Jacobsen, L.B.; Pedersen, S.M.; Tavella, E. Socioeconomic Impact of Widespread Adoption of Precision Farming and Controlled Traffic Systems in Denmark. *Precis. Agric.* **2012**, *13*, 661–677. [[CrossRef](#)]
4. 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](#)]
5. Lu, J.; Tan, L.; Jiang, H. Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. *Agriculture* **2021**, *11*, 707. [[CrossRef](#)]
6. Mondal, D.; Kole, D.K.; Roy, K. Gradation of Yellow Mosaic Virus Disease of Okra and Bitter Gourd Based on Entropy Based Binning and Naive Bayes Classifier after Identification of Leaves. *Comput. Electron. Agric.* **2017**, *142*, 485–493. [[CrossRef](#)]
7. Thaiyalnayaki, K.; Joseph, C. Classification of Plant Disease Using SVM and Deep Learning. *Mater. Today Proc.* **2021**, *47*, 468–470. [[CrossRef](#)]
8. Tian, K.; Li, J.; Zeng, J.; Evans, A.; Zhang, L. Segmentation of Tomato Leaf Images Based on Adaptive Clustering Number of K-Means Algorithm. *Comput. Electron. Agric.* **2019**, *165*, 104962. [[CrossRef](#)]
9. Aravind, K.R.; Raja, P.; Mukesh, K.V.; Anirudh, R.; Ashiwin, R.; Szczepanski, C. Disease Classification in Maize Crop Using Bag of Features and Multiclass Support Vector Machine. In Proceedings of the 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 19–20 January 2018; pp. 1191–1196.
10. Hossain, E.; Hossain, M.F.; Rahaman, M.A. A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier. In Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 7–9 February 2019; pp. 1–6.
11. Kahar, M.A.; Mutalib, S.; Abdul-Rahman, S. Early Detection and Classification of Paddy Diseases with Neural Networks and Fuzzy Logic. In Proceedings of the 17th International Conference Computational and Mathematical Methods in Science and Engineering, CMMSE 2017, Rota, Spain, 4–8 July 2017; pp. 248–257.
12. Abade, A.; Ferreira, P.A.; de Barros Vidal, F. Plant Diseases Recognition on Images Using Convolutional Neural Networks: A Systematic Review. *Comput. Electron. Agric.* **2021**, *185*, 106125. [[CrossRef](#)]
13. Nie, X.; Wang, L.; Ding, H.; Xu, M. Strawberry Verticillium Wilt Detection Network Based on Multi-Task Learning and Attention. *IEEE Access* **2019**, *7*, 170003–170011. [[CrossRef](#)]
14. Rangarajan, A.K.; Purushothaman, R.; Pérez-Ruiz, M. Disease Classification in Aubergine with Local Symptomatic Region Using Deep Learning Models. *Biosyst. Eng.* **2021**, *209*, 139–153. [[CrossRef](#)]
15. Waheed, A.; Goyal, M.; Gupta, D.; Khanna, A.; Hassani, A.E.; Pandey, H.M. An Optimized Dense Convolutional Neural Network Model for Disease Recognition and Classification in Corn Leaf. *Comput. Electron. Agric.* **2020**, *175*, 105456. [[CrossRef](#)]
16. Ramcharan, A.; Baranowski, K.; McCloskey, P.; Ahmed, B.; Legg, J.; Hughes, D.P. Deep Learning for Image-Based Cassava Disease Detection. *Front. Plant Sci.* **2017**, *8*, 1852. [[CrossRef](#)]
17. Haider, W.; Rehman, A.-U.; Durrani, N.M.; Rehman, S.U. A Generic Approach for Wheat Disease Classification and Verification Using Expert Opinion for Knowledge-Based Decisions. *IEEE Access* **2021**, *9*, 31104–31129. [[CrossRef](#)]
18. Zeng, T.; Li, C.; Zhang, B.; Wang, R.; Fu, W.; Wang, J.; Zhang, X. Rubber Leaf Disease Recognition Based on Improved Deep Convolutional Neural Networks with a Cross-Scale Attention Mechanism. *Front. Plant Sci.* **2022**, *13*, 829479. [[CrossRef](#)]

19. Mi, Z.; Zhang, X.; Su, J.; Han, D.; Su, B. Wheat Stripe Rust Grading by Deep Learning with Attention Mechanism and Images from Mobile Devices. *Front. Plant Sci.* **2020**, *11*, 558126. [[CrossRef](#)] [[PubMed](#)]
20. Zhang, D.; Wang, D.; Gu, C.; Jin, N.; Zhao, H.; Chen, G.; Liang, H.; Liang, D. Using Neural Network to Identify the Severity of Wheat Fusarium Head Blight in the Field Environment. *Remote Sens.* **2019**, *11*, 2375. [[CrossRef](#)]
21. Fang, T.; Chen, P.; Zhang, J.; Wang, B. Crop Leaf Disease Grade Identification Based on an Improved Convolutional Neural Network. *J. Electron. Imaging* **2020**, *29*, 013004. [[CrossRef](#)]
22. Wang, C.; Du, P.; Wu, H.; Li, J.; Zhao, C.; Zhu, H. A Cucumber Leaf Disease Severity Classification Method Based on the Fusion of DeepLabV3+ and U-Net. *Comput. Electron. Agric.* **2021**, *189*, 106373. [[CrossRef](#)]
23. Buda, M.; Maki, A.; Mazurowski, M.A. A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks. *Neural Netw.* **2018**, *106*, 249–259. [[CrossRef](#)]
24. Dyrmann, M.; Karstoft, H.; Midtiby, H.S. Plant Species Classification Using Deep Convolutional Neural Network. *Biosyst. Eng.* **2016**, *151*, 72–80. [[CrossRef](#)]
25. Yu, F.; Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions. *arXiv* **2016**, arXiv:1511.07122.
26. Shi, W.; Jiang, F.; Zhao, D. Single Image Super-Resolution with Dilated Convolution Based Multi-Scale Information Learning Inception Module. In Proceedings of the 2017 24th IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: New York, NY, USA, 2017; pp. 977–981.
27. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
28. Tang, Z.; Yang, J.; Li, Z.; Qi, F. Grape Disease Image Classification Based on Lightweight Convolution Neural Networks and Channelwise Attention. *Comput. Electron. Agric.* **2020**, *178*, 105735. [[CrossRef](#)]
29. Chen, J.; Wang, W.; Zhang, D.; Zeb, A.; Nanehkaran, Y.A. Attention Embedded Lightweight Network for Maize Disease Recognition. *Plant Pathol.* **2021**, *70*, 630–642. [[CrossRef](#)]
30. Xingyan, L.; Dan, Q. Joint Bottleneck Feature and Attention Model for Speech Recognition. In Proceedings of the 2018 International Conference on Mathematics and Artificial Intelligence (ICMAI 2018), New York, NY, USA, 20–22 April 2018; Assoc Computing Machinery: New York, NY, USA, 2018; pp. 46–50.
31. Mi, H.; Sankaran, B.; Wang, Z.; Ittycheriah, A. Coverage Embedding Models for Neural Machine Translation. *arXiv* **2016**, arXiv:1605.03148.
32. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/Cvf Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; IEEE: New York, NY, USA, 2018; pp. 7132–7141.
33. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 21–26 July 2016; IEEE: New York, NY, USA, 2017; pp. 2261–2269.