

Article

Identification Method of Corn Leaf Disease Based on Improved Mobilenetv3 Model

Chunguang Bi ^{1,2}, Suzhen Xu ², Nan Hu ², Shuo Zhang ², Zhenyi Zhu ² and Helong Yu ^{1,2,*}¹ Institute for the Smart Agriculture, Jilin Agricultural University, Changchun 130118, China² College of Information Technology, Jilin Agricultural University, Changchun 130118, China

* Correspondence: yuhelong@jlau.edu.cn; Tel.: +86-135-0082-8956

Abstract: Corn is one of the main food crops in China, and its area ranks in the top three in the world. However, the corn leaf disease has seriously affected the yield and quality of corn. To quickly and accurately identify corn leaf diseases, taking timely and effective treatment to reduce the loss of corn yield. We proposed identifying corn leaf diseases using the Mobilenetv3 (CD-Mobilenetv3) model. Based on the Mobilenetv3 model, we replaced the model's cross-entropy loss function with a bias loss function to improve accuracy. Replaced the model's squeeze and excitation (SE) module with the efficient channel attention (ECA) module to reduce parameters. Introduced the cross-layer connections between Mobile modules to utilize features synthetically. Then we Introduced the dilated convolutions in the model to increase the receptive field. We integrated a hybrid open-source corn leaf disease dataset (CLDD). The test results on CLDD showed the accuracy reached 98.23%, the precision reached 98.26%, the recall reached 98.26%, and the F1 score reached 98.26%. The test results are improved compared to the classic deep learning (DL) models ResNet50, ResNet101, ShuffleNet_x2, VGG16, SqueezeNet, InceptionNetv3, etc. The loss value was 0.0285, and the parameters were lower than most contrasting models. The experimental results verified the validity of the CD-Mobilenetv3 model in the identification of corn leaf diseases. It provides adequate technical support for the timely control of corn leaf diseases.

Keywords: CNN; disease; corn; image processing; identification

Citation: Bi, C.; Xu, S.; Hu, N.; Zhang, S.; Zhu, Z.; Yu, H. Identification Method of Corn Leaf Disease Based on Improved Mobilenetv3 Model. *Agronomy* **2023**, *13*, 300. <https://doi.org/10.3390/agronomy13020300>

Academic Editors: Wen-Hao Su and Zhou Zhang

Received: 31 October 2022

Revised: 9 January 2023

Accepted: 11 January 2023

Published: 18 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

Corn is the most productive grain crop in the world and one of the major grain crops in China. China's corn industry is developing rapidly; the sown area reached 41.26 million hm² in 2020, creating a high yield record of 24,948.75 kg/hm² [1]. Various leaf diseases occur in corn during planting, affecting corn's yield and quality. Therefore, it is essential to efficiently and accurately identify corn leaf diseases. Solve the following problems intense subjectivity and low efficiency in traditional manual identification of diseases. Some researchers, such as Dang [2], Xiong [3], Su [4], Yuan [5], etc., used image processing technology based on machine learning to diagnose and identify types of crop diseases. Traditional machine learning [6–10] requires manual extraction of disease features, and it is not only time-consuming and labor-intensive but also a low generalization. The emergence of deep learning (DL) [11–15] provides a new method for plant disease identification. DL is also used in many fields [16–20]. Convolutional neural network(CNN) [21–25] is one of the most common DL methods. In image classification, due to the large amount of data to be processed, it isn't easy to keep the original features of images in digitization. CNN has the ability of representation learning, which can automatically extract features from images of different scales. CNN has a strong ability to extract image features and high accuracy. It has been widely used in agriculture in recent years. For example, In 2020, Xu et al. [26] proposed a CNN model based on the VGG-16 model to realize the image recognition of corn diseases under the complex field background of small data samples.

This model uses transfer learning to recognize corn disease images, which can improve the model's convergence speed and recognition ability. The accuracy is 95.33%. Ren et al. [27] constructed a VGG network model based on deconvolution guidance for plant leaf disease identification and disease spot segmentation. Anagnostis et al. [28] used a self-built walnut leaf dataset, they proposed a CNN-based method for identifying anthracnose in walnut leaves, and the accuracy ranges from 92.4% to 98.7%. Maeda-Gutiérrez et al. [29] used models such as AlexNet, GoogleNet, and InceptionV3 to identify nine different types of tomato diseases, and the accuracy is 99.72%. In 2021, Bao et al. [30] proposed an improved CNN model to identify corn leaf diseases, it improved the model's stability, and the accuracy was 95.74%. Hassan et al. [31] proposed two methods, shallow VGG with RF and shallow VGG with Xgboost, to identify diseases, and the experiments achieved good results on corn, potato, and tomato. In 2022, Wang et al. [32] proposed a corn disease identification model (AT-AlexNet) based on AlexNet, constructed a new downsampling attention module, and introduced the Mish activation function, and the average accuracy was 99.35%. The above studies have achieved good results, but the model parameters are large, and the running time is long. These are unsuitable for mobile terminal development and not convenient for farmers. It isn't easy to expand other applications.

Some scholars have also proposed lightweight networks to run CNN models on mobile and embedded devices. It has the advantages of fewer parameters and smaller model sizes, such as MobileNetv1 [33], MobileNetv2 [34], ShuffleNetv1 [35], and ShuffleNetv2 [36], etc. In 2020, Mi et al. [37] proposed a new DL network (C-DenseNet) based on the DenseNet web (C-DenseNet), Embedded the convolutional block attention module (CBAM) into the densely connected convolutional net DenseNet, and the accuracy is 97.99%. Chao et al. [38] proposed a DL network model XDNet based on DenseNet and Xception, it identified five apple leaf diseases, and the accuracy was 98.82%. In 2021, Liu et al. [39] used the PlantVillage public dataset as the experimental data. They studied a lightweight network based on the improved SqueezeNet model, which identified many leaf diseases. Sun et al. [40] embedded a lightweight coordinate attention mechanism in the model MobileNetV2; it reduced the model's parameters and improved its accuracy. In 2022, Li et al. [41] proposed a method based on ResNet lightweight residual network (Scale Down ResNet) to identify plant leaf diseases. It reduced network parameters and computational complexity and maintained a low identification error rate. Zeng et al. [42] proposed a group multi-scale attention network to identify rubber leaf disease images (GMA-Net), experiments on the constructed rubber leaf disease dataset, and the PlantVillage public dataset. The accuracy of the model is 98.06% and 99.43%, respectively. Eunice et al. [43] used a CNN-based pre-training model to identify plant diseases effectively. They were fine-tuning the hyperparameters of the popular pre-training model and conducting experiments on the PlantVillage data set. Experiments showed that the classification accuracy of DenseNet-121 is 99.81%, which was superior to the most advanced model. Some researchers began using mobile phones to identify diseases to make it convenient for farmers. In 2019, Liu et al. [44], based on MobileNet and Inception V3 network combined with transfer learning, proposed two crop disease classification models to detect plant leaf disease on mobile phones. In 2020, Yu et al. [45], based on DL and combined with transfer learning, proposed a CDCNNv2 model based on a residual network(ResNet 50) to detect crop diseases and pests. They designed an Android-based crop pest identification APP. In 2021, Fan et al. [46] proposed a detection system for grape leaf disease based on transfer learning and improved VGG16 on Android mobile phones. It obtained images by taking photos or acquiring the local gallery, clicking the identification button, and the system output the results of grape diseases.

Based on previous research, to reduce the complexity and parameters of the model and improve the accuracy of the model. We used Mobilenetv3 as the basic model of our research, and it is a lightweight CNN. Mobilenetv3 [47] uses the cross-entropy loss function to measure the difference between two probability distributions and between the learned and actual distributions. Mobilenetv3 has the squeeze and excitation (SE) [48] module. The SE module adds an attention mechanism in the channel dimension. The critical operations

are squeezes and exceptions. It obtains the importance of each channel of the feature map. It Uses this importance to assign a weight value to each feature, and the neural network can focus on some feature channels. We have improved the model based on Mobilenetv3. The main innovations of this paper are as follows:

- In the Mobilenetv3 model, we have replaced the cross-entropy loss function with the Bias Loss function, which can reduce problems the random predictions can cause during optimization;
- We have replaced the SE module in the model with the efficient channel attention (ECA) module, which can reduce the parameters of the model and improve its accuracy of the model;
- The cross-layer connections are introduced between Mobile modules, introducing the shallow features into deep layers. It enables effectively comprehensive utilization of local features and deep features, and the extracted features are more extensive;
- The dilated convolutions are introduced in the first and last convolution of the model, increasing the receptive field so that convolution can extract more information;
- The experimental results can verify the validity of the improved model, and it can reduce the model's parameters and shorten the model's running time, improving its accuracy.

2. Materials and Methods

2.1. Data Sources

Our research used mixed open-source corn leaf disease datasets (CLDD). The CLDD contained different public datasets on the Internet, including PlantVillage [49], AI challenger 2018, PlantifyDr, and PlantDoc [50] datasets, etc. Most of these data were from the research stations of field trials in many countries. Researchers usually used standard digital cameras and smartphones to take images of plant leaves. Our research integrated most of the data related to corn leaf diseases. The CLDD has 25,167 images, including four kinds of corn leaf diseases (northern leaf blight, dwarf mosaic virus, rust, gray spot) and healthy leaves. This CLDD contained 6765 images of corn north leaf blight, 931 images of corn dwarf mosaic virus, 6126 images of corn rust disease, 3588 images of corn grey spot disease, and 7757 images of healthy corn leaves. Figure 1 shows partial examples of the original CLDD.

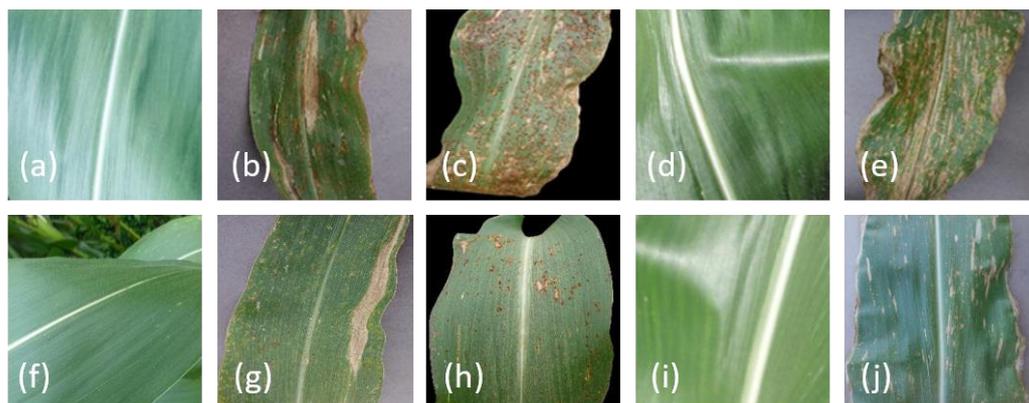


Figure 1. Partial examples of different levels of disease in the original dataset. Note: (a–e) Severe corn leaf disease; (f–j) Mild corn leaf disease; (a,f) Healthy; (b,g) Northern leaf blight; (c,h) Common rust; (d,i) Dwarf mosaic virus; (e,j) Gray leaf spot.

2.2. Data Preprocessing

Our research set the image size to $448 \times 448 \times 3$. To balance the samples and reduce the overfitting phenomenon of the network. Our research preprocessed the dataset and enhanced some datasets. The enhancement method used a variety of ways, such as adding noise, mirror transformation, horizontal flipping, and clipping. The specific operations are as follows:

- By flipping, adding noise, rotation, blurring, etc., dwarf mosaic virus image data has increased six-fold, a total of 5586 images after enhancement;
- By adding noise and flipping, Gray spot image data doubled, totaling 7176 images after enhancement. A total of 33,409 images of the enhanced corn leaf disease dataset.

Our research randomly selected 80% of the dataset as the training set and 20% of the dataset used as the test set. The training and test sets are 26,727 and 6682 images, respectively. Figure 2 shows examples of the augmented datasets, and Table 1 shows the number of datasets before and after the enhancement process.

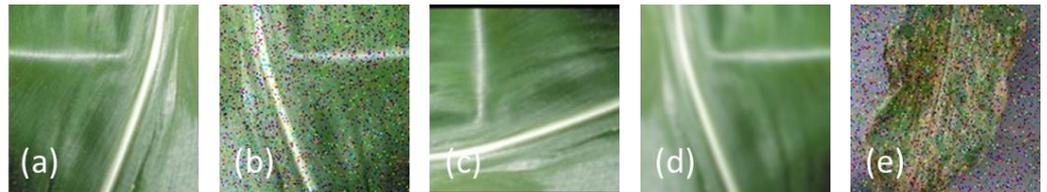


Figure 2. Partial examples of the enhanced dataset. Note: (a) Fli; (b) Noise; (c) R90; (d) Blur; (e) Noise.

Table 1. Original and enhanced images of corn leaf disease datasets(CLDD).

Disease Class	Original Image	Enhanced Image	Sample Label
Common rust	6125	6125	1
Gray leaf spot	3588	7176	2
Dwarf mosaic virus	931	5586	3
Healthy	7757	7757	4
Northern leaf blight	6765	6765	5
Total	25,166	33,409	5

2.3. Experiment Method

2.3.1. Mobilenetv3 Network

Mobilenetv3 [33,47] obtained parameters by NAS (network architecture search) search. It has two versions, large and small, suitable for different scenarios. It inherits the depthwise separable convolution of Mobilenetv1 and the residual structure with the linear bottleneck of Mobilenetv2. Mobilenetv3 uses the NetAdapt algorithm to obtain the optimal number of convolution kernels and channels. It introduces the SE channel attention structure based on MobileNetV2 and modifies the MobileNetV2 back-end output. Mobilenetv3 uses a new activation function, h-swish (x), instead of Relu6. It uses Relu6 (x + 3)/6 to simulate sigmoid in the SE module. Figure 3 shows the network model structure of Mobilenetv3, divided into three parts: The first part consists of 1 convolutional layer, extracting features by 3 × 3 convolution; the second part has multiple convolutional layers; due to different levels and parameters, divided into large and small version, the number of small is 13 and large is 15; the third part is to reduce parameters and calculation, it advances the Avg Pooling, replacing the entire connection with two 1 × 1 convolutional layers, and finally outputting the category.

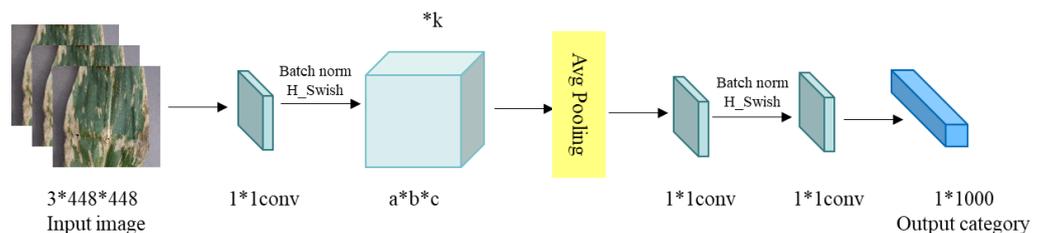


Figure 3. Structure diagram of Mobilenetv3 network model. Note: “k” represents the number of Mobile modules, “a*b*c” means the channel number and size of the Mobile module, including 16*224*224, 24*112*112, 40*56*56, 80*28*28, 112*28*28 and 160*14*14.

2.3.2. Dilated Convolution

The dilated convolution [51] aims at the problem of image semantic segmentation, and it is an idea that downsampling reduces image resolution and misses information. The convolution receptive field increases when the parameter quantity is unchanged. Dilated convolution introduces a new parameter called “dilation rate” to the convolutional layer, which defines the spacing of each value when the convolution kernel processes the data. It expands the receptive field by adding holes under the same parameters and calculation. It allows the original 3×3 convolution kernel to have a 5×5 (dilated rate = 2) or larger receptive field, so there is no need to downsample. Different receptive fields bring multi-scale information when multiple dilated convolution kernels with varying dilation rates are stacked. Dilated convolution can expand the receptive field without losing resolution and keep the relative spatial position of the pixels unchanged.

2.3.3. Bias Loss Function

Bias Loss [52] is a dynamic scaling cross-entropy loss, and its scale decays as the variance of the data points decreases. The Bias Loss function helps to focus learning on samples that can provide many unique features. It reduces problems that random prediction causes during optimization. Bias Loss defines as:

$$L_{bias} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k z(v_i) y_{ij} \log f_j(X_i; \theta), \quad (1)$$

$$z(v_i) = \exp(v_i * \alpha) - \beta, \quad (2)$$

Letting $X \in R^{C \times h \times w}$ be the feature space. c is the several input channels. h, w is the height and width of the input data. $Y = \{1, \dots, k\}$ is the label space. k is the number of classes. A standard scenario has a dataset $D = (x_i, y_i)_{i=1}^N$, each $(x_i, y_i) \in X \times Y$ neural network $f(X; \theta)$. In the formula, α and β are adjustable parameters. v is the scaling variance of the convolution layer output. The non-linear scaling function $z(v_i)$ generates a bias between low and high-variance data points.

2.3.4. ECA Module

ECA module [53] is an efficient channel attention mechanism. This module involves a few parameters and brings significant performance gains. The ECA module includes an average pooling layer, a 1×1 convolution, and a sigmoid activation function. ECA module based on the SE module replacing multilayer perceptron (MLP) module with one-dimensional convolution, avoiding dimension reduction and effectively realizing cross-channel interaction. ECA module realizes information interaction between channels through $1D$ convolution with a kernel size of k , i.e.,

$$\omega = \sigma(C1D_k(y)), \quad (3)$$

where σ is a Sigmoid function, $1D$ indicates $1D$ convolution. k is kernel size. y is the aggregated feature without dimension reduction. There may exist a mapping ϕ between k and C . Channel dimension C usually is set to a power of 2. However, the relations characterized by linear function are too limited. Therefore, we introduce a possible solution by extending the linear function $\phi(k) = \gamma * k - b$ to a non-linear one, i.e.,

$$C = \phi(k) = 2^{\gamma * k - b}, \quad (4)$$

where C is the channel dimension, and b is constant. Then, given channel dimension C , kernel size k can be adaptively determined by

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}}, \quad (5)$$

where $|t|_{odd}$ indicates the nearest odd number of t . By mapping ψ , High dimensional channels have more extended range interactions, while low dimensional channels experience shorter range interactions using nonlinear mapping.

Our research replaced the cross-entropy loss function in the model with the Bias Loss function, and it reduced problems caused by random predictions during optimization. Introduced dilated convolution increased the receptive field and let convolution extract more information. Replaced the SE module in the model with the ECA module and effectively reduced the amount of parameter calculation. Introduced cross-layer connections between Mobile modules, introduced shallow features into deep layers, and effectively utilized local and in-depth features. Our research constructed an identification Mobilenetv3 model of corn leaf diseases (CD-Mobilenetv3). Figure 4 shows the network model structure of CD-Mobilenetv3. The specific implementation steps are as follows:

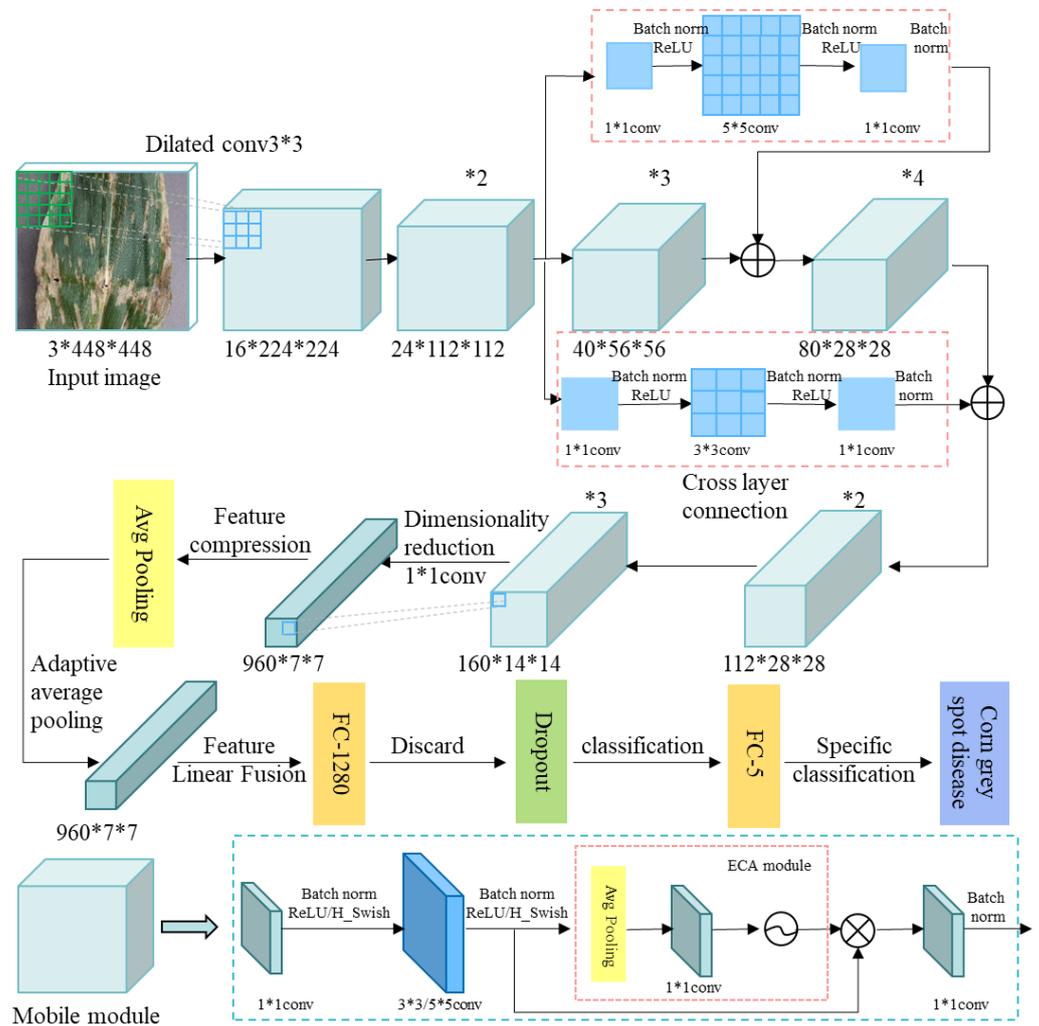


Figure 4. Structure diagram of CD-Mobilenetv3 network model. Note: “ σ ” represents Sigmoid function, “ \otimes ” represents the weighting operation of the matrix, “ \oplus ” represents the addition operation of the convolutional features vector.

Introduced dilated convolution: we introduced dilated convolution in the 3*3 convolution of the first Mobile module and the 5*5 convolution of the last Mobile module of the model. Dilated convolution increased the receptive field so that convolution extracted more information. First, we input a corn leaf disease image with a size of 448*448*3. It acquired features through a 3*3 dilated convolution. The extracted image size became 16*224*224,

and its number of channels increased to 16. The picture size became half of the original. Dilated convolution improved model accuracy and reduced loss value.

Replaced the SE module in the Mobile module with the ECA module: The CD-MobileNetv3 replaced the SE module in the Mobile module with the ECA module, improved the model's accuracy, and reduced the model's parameters. The ECA module includes an average pooling layer, a 1*1 convolution, and a sigmoid activation function. The process of the improved Mobile module was through a 1*1 convolution, Batch norm, and ReLU/H_Swish activation function. The Batch norm accelerated the network convergence speed, and the activation function reduced the amount of calculation. Then it went through a 3*3/5*5 convolution, Batch norm, and ReLU/H_Swish activation function. It calculates the feature weight through the ECA module and multiplies the original feature and the feature weight to obtain a weighted feature set. Then input to 1*1 convolution for channel dimension reduction, output to the following Mobile module after passing through the Batch norm.

Introduced cross-layer connections: To reduce the loss of features in the transfer process, we adopted cross-layer connections between mobile modules to integrate the characteristics of different layers. When the number of channels and the Mobile module with the size of 24*112*112 is output, made two cross-layer connections: The first cross-layer connection went through 1*1 convolution, Batch norm, and ReLU functions. It passed the 5*5 convolution, Batch norm, and ReLU functions, output after 1*1 convolution and Batch norm. These operations increased the number of channels and reduced dimensions. It performed a feature sum operation with the Mobile module, which has many channels and sizes of 40*56*56, and passed the feature result to the following Mobile module. The second cross-layer connection replaced the 5*5 convolution in the first step with a 3*3 convolution. The rest of the steps were similar. These operations introduced the features of the shallow layer into the deep layer, effectively utilized local and in-depth features, and extracted more comprehensive features.

Feature fusion: Stacking 15 Mobile modules formed a deep network and extracted more feature information. It underwent 960*1*1 convolutions for dimensionality reduction, then output a feature set of 960*7*7. It performed its average pooling and feature compression through a pooling layer. After a fully connected layer with 1280 categories, linearly fused the features extracted from the convolution were.

Specific classification: To increase the model's generalization and reduce the model's overfitting in the model, we added a random dropout behind a fully connected layer with 1280 categories. Then, we added a fully connected layer with five classes, whose function is non-linear change and specific classification. Finally, output the identified results.

2.3.5. Evaluation Index

The evaluation index used the accuracy rate to calculate the proportion of correct classification. However, due to sample imbalance, it combines precision, recall, and F1 score to evaluate the model comprehensively. The higher the precision, the stronger the model's ability to distinguish negative samples. The higher the recall, the stronger the model's ability to identify positive samples. The higher the F1 score, the more robust the model. The calculation methods for each indicator are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

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

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

$$F1 = \frac{2Precision \times Recall}{Precision + Recall} \quad (9)$$

In the formula, TP is the number of actual positive images of corn leaf disease; FP is the number of false positive images of corn leaf disease; TN is the number of actual negative images of corn leaf disease; FN is the number of false negative images of corn leaf disease.

3. Results

3.1. Experimental Environment and Parameter Settings

The test environment is Ubuntu 18.04 64-bit operating system, and the python version is 3.7. The DL framework is Python 1.8.1. Hardware environment: The processor on the computer is Intel(R) Xeon(R) Gold 6246R CPU @ 3.40 GHz (64 CPUs), ~3.4 GHz. The running memory is 128 GB, the graphics card is NVIDIA Quadro RTX 8000, and the video memory (VRAM) is 48 GB.

The input image size was 448*448*3. The number of training iterations epoch was 100. The batch size when training disease classification was 64. The learning rate of the model was 0.01. To reduce problems such as overfitting, the training model used the AdamW optimizer, and weight_decay was a default.

3.2. Ablation Experiment

Based on the Mobilenetv3 model, our research introduced the Bias Loss function, dilated convolution, and ECA module and introduced cross-layer connections between Mobile modules. It experimented on the CLDD to verify the performance of the CD-Mobilenetv3 model. We got the experimental results through the test set. Table 2 shows the results of the ablation test.

Table 2. Ablation test results of CLDD.

Model	LossValue	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Params
Mobilenetv3-large	0.1314	95.54	95.71	95.69	95.61	5.48 M
Mobilenetv3-large-bia	0.0568	96.72	96.82	96.76	96.78	5.48 M
Mobilenetv3-large-bia-eca	0.0540	96.38	96.47	96.45	96.45	3.97 M
Mobilenetv3-large-bia-eca-skip	0.0400	97.47	97.51	97.51	97.51	4.36 M
Mobilenetv3-large-bia-eca-skip-digconv (CD-Mobilenetv3)	0.0285	98.23	98.26	98.26	98.26	4.36 M

Table 2 shows the results of the ablation experiments. The accuracy of the CD-Mobilenetv3 model on the corn leaf disease dataset reached 98.23%. Based on the Mobilenetv3 model, we replaced the cross-entropy loss function in the model with the Bias Loss function. Its accuracy was 1.18% higher than the basic model on the data set. It showed that the Bias Loss function focused the learning of the model on samples that provided a large number of unique features, so that effectively extracted more features. Then we replaced the SE module in the model with the ECA module. The accuracy of the experimental results was 0.84% higher than the basic model and reduced its parameters by 27.56%. It showed that the ECA module improved the model accuracy and effectively reduced the amount of parameter calculation. Introduced cross-layer connections, its accuracy rate was 1.93% higher than Mobilenetv3 and reduced its parameters by 20.44%. It showed that cross-layer connections effectively integrated local and in-depth features, and the extracted features were more comprehensive. Introduced dilated convolution in the first and last convolution of the model, the accuracy of the CD-Mobilenetv3 model on the dataset was 2.69% higher than the Mobilenetv3. It showed that after dilated convolution expanded the receptive field, it better extracted the semantic information of leaves and lesions. In terms of the loss value, the CD-Mobilenetv3 model was 0.1029 lower on the dataset than the basic model. In terms of precision, the CD-Mobilenetv3 model was 2.55% more promoted than the basic model. In terms of recall, the CD-Mobilenetv3 model improved by 2.57% on the dataset compared to the basic model. In the aspect of the F1 score, the CD-Mobilenetv3 model was 2.65% higher than the basic model in a dataset, and its parameters were reduced by

20.44%. It showed that the results of the ablation experiment verified the effectiveness of the CD-Mobilenetv3 model.

Figure 5 shows the convergence of the accuracy and loss value of the improved network model on the test set within 100 epochs. Figure 5a shows that with the increase in iteration times, the addition of other improved models gradually slowed to around 20 epochs. It started to converge at about 40 epochs, while the CD-Mobilenetv3 curve rose the fastest and grew after 20 epochs. It reached the highest point when training to 40 epochs, and the training accuracy of the network remained stable and ranked at the top. It showed that introducing the ECA module reduced the model's parameters and improved computing efficiency. Introduced cross-layer connections and dilated convolution better-extracted feature information and significantly improved the model's accuracy. As seen from Figure 5b, the loss value of the improved network model gradually decreased under the training of data samples, and the CD-Mobilenetv3 curve decreased the fastest. After 40 epochs of training, the loss value of the network reached the lowest value, which approached 0.0285. It showed that the model has good robustness, which reflected that the Bias Loss function reduced the problems caused by random prediction in the optimization process. Therefore, the improvement effect of the CD-Mobilenetv3 model was ideal, and it was more suitable for classifying corn leaf diseases.

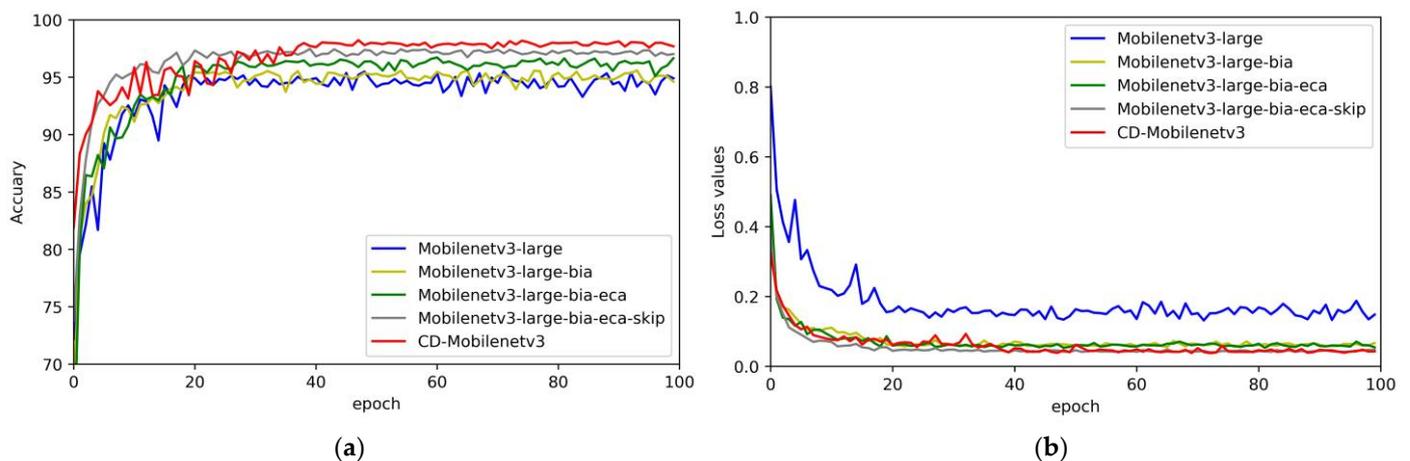


Figure 5. Accuracy and loss of improved network model. (a) Accuracy; (b) Loss value.

3.3. Comparative Experiment

To further verify the effectiveness of the CD-Mobilenetv3 model proposed in this paper. Compared heavyweight network models ResNet50, ResNet101, VGG16, and lightweight network models ShuffleNet_x2, SqueezeNet, InceptionNetv3, Mobilenetv3-small and Mobilenetv3-large for experiments. The comparison models were all tested on the original framework of the model and parameter settings. The experimental results were all obtained on the test set. Table 3 shows the comparative test results.

Table 3 shows the results of the comparative experiments. Compared with the heavy-weight network model and the lightweight network model, the CD-Mobilenetv3 model has a significant improvement in accuracy. In the experimental CLDD, the accuracy of the CD-Mobilenetv3 model proposed by the study reached 98.23%. Compared with ResNet50, ResNet101, ShuffleNet_x2, VGG16, SqueezeNet, InceptionNetv3, Mobilenetv3-small and Mobilenetv3-large, the improvements were 2.00%, 3.82%, 6.39%, 1.62%, 5.89%, 3.28%, 3.51% and 2.69% respectively. In terms of loss value, compared with other models, the CD-Mobilenetv3 model reduced by 0.0203~0.1854. In terms of precision, the CD-Mobilenetv3 model improved by 1.59% to 6.24% compared to other models on the dataset. In terms of recall, the CD-Mobilenetv3 model improved by 1.61% to 6.42% compared to other models on the dataset. On the dataset, the F1-score of the CD-Mobilenetv3 model was 1.61%–6.67% higher than other models, and the parameters were lower than most of the comparison

models 1.12–134 M. The CD-Mobilenetv3 model was more accurate than the experimental results, and its various indicators are also more suitable. It has less parameter calculation and a shorter running time. Therefore, the CD-Mobilenet v3 model was more suitable for developing and applying mobile software.

Table 3. Comparative experimental results of CLDD.

Model	Loss Value	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Params
ResNet50	0.1151	96.23	96.38	96.36	96.30	25.56 M
ResNet101	0.1642	94.42	93.99	93.68	93.54	44.55 M
ShuffleNet_x2	0.2139	91.84	92.70	92.19	92.06	7.39 M
VGG16	0.0488	96.61	96.67	96.65	96.65	138.36 M
Mobilenetv3-small	0.1634	94.72	95.00	94.84	94.87	2.54 M
SqueezeNet	0.2129	92.34	92.02	91.84	91.59	1.24 M
Mobilenetv3-large	0.1314	95.54	95.71	95.69	95.61	5.48 M
InceptionNetv3	0.1533	94.96	94.34	94.18	94.02	27.16 M
CD-Mobilenetv3	0.0285	98.23	98.26	98.26	98.26	4.36 M

Figure 6 shows the convergence of the accuracy and loss values of the network models compared to the test set within 100 epochs. Figure 6a shows that adding other DL models' accuracy gradually slowed down around 20 epochs with the increase in the number of iterations. It began to converge around 40 epochs, and the accuracy of the overall network model was not high. After 20 epochs of training, the accuracy of the CD-Mobilenetv3 model still improved gradually under the training of data samples. The CD-Mobilenetv3 model began to converge after about 40 epochs of training, and the training accuracy of the network approached 98.23%. Figure 6b shows that with the increase in iteration times, the loss value of other comparable models decreased slowly. The loss value of the CD-Mobilenetv3 model dropped fastest and reached a stable point in 40 epochs. The loss value of the network approached 0.0285. From the perspective of the convergence of the loss value, the training effect of the CD-Mobilenetv3 model was ideal.

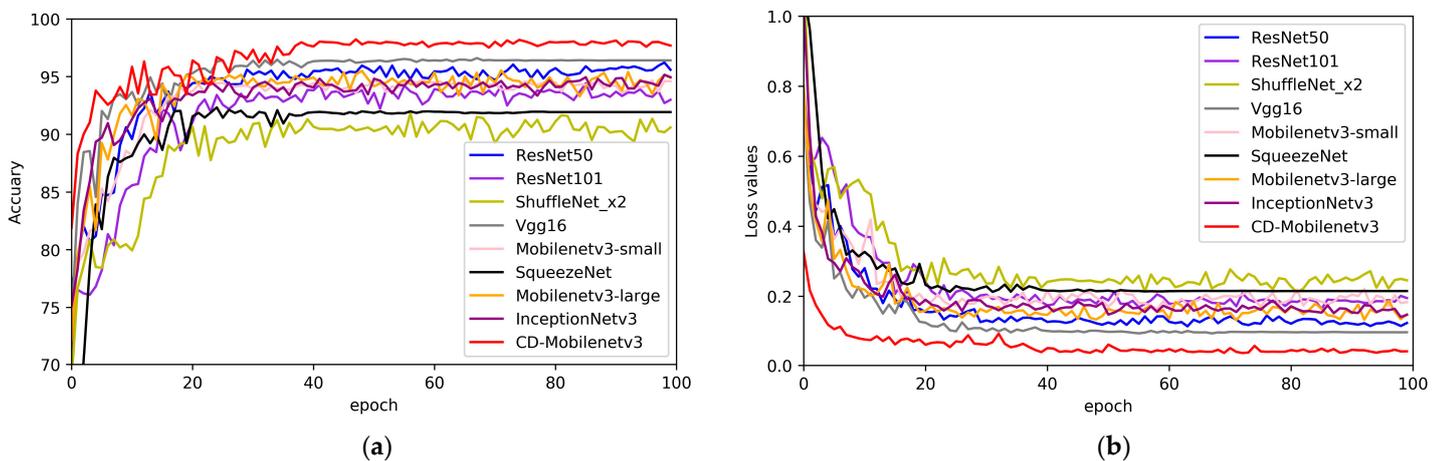


Figure 6. Accuracy and loss of comparative network model. (a) Accuracy; (b) Loss value.

3.4. Feature Attention Visualization

Gradient-weighted class activation mapping (Grad-CAM) [54] is a specific positioning technology category that can generate visual interpretation for any CNN-based network. It uses the gradient of any target concept flowing into the final convolution layer to generate a rough location map, highlighting the critical areas in the image used to predict concepts. After introducing the attention mechanism into the basic model, further verified the impact of the CD-Mobilenetv3 model on model performance improvement. Our research selected some examples of the CLDD, then used Grad-CAM to visualize the feature of the CD-

Mobilenetv3 model. The feature attention visualization heatmap showed that the CD-Mobilenetv3 model did convolution or other operations. It was the process of extracting image features. Feature heatmap highlighted its detailed features, and the feature map output by the network focused more on the feature part of the image.

It can be seen more intuitively from Figure 7, where the corn leaf disease was deep and bright, and features with different colors showed the attention of different regions. The redder the color, the higher the attention of the area. It showed that when introduced the attention mechanism, the CD-Mobilenetv3 model paid more attention to the regions with more apparent features in the process of classification and identification in feature extraction and representation. It reduced problems such as background area interference in the disease identification process to improve the network model's performance.

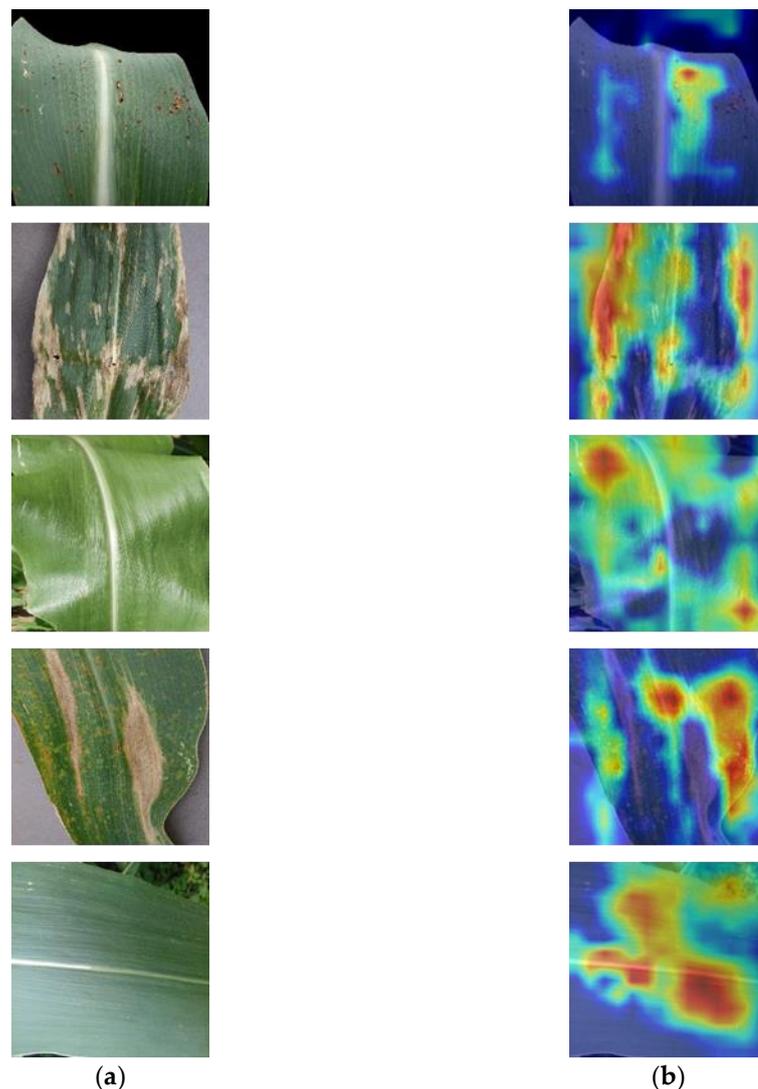


Figure 7. Visual diagram of features attention. (a) Original image; (b) Feature image.

3.5. Confusion Matrix

Analyzing the confusion matrix of recognition results presented in Figure 8 showed that by using the CD-Mobilenetv3 model, from the error between actual and predicted labels, labels 2 and 5 have a high misclassification rate. Label 1 was mistakenly divided into label 2 with thirteen images, label 1 was mistakenly divided into label 3 with five images, label 1 was mistakenly divided into label 4 with one image, and label 1 was mistakenly divided into label 5 with seven images. Label 2 was mistakenly divided into label 3

with three images, label 2 was mistakenly divided into label 4 with 1 image, and label 2 was mistakenly divided into label 5 with forty-six images. Label 3 has no mismarking. Label 4 was mistakenly divided into label 3, with two images total. Label 5 was mistakenly divided into label 1 with six images. Label 5 was mistakenly divided into label 2 with thirty-one images, label 5 was mistakenly divided into label 3 with one image, and label 5 was mistakenly divided into label 4 with two images. These results further verified the effectiveness of the attention mechanism in improving model recognition performance. Further analysis of the results showed that the disease features of the misclassified images were very similar, so the accurate recognition of image features has a certain influence on the accurate recognition of disease.

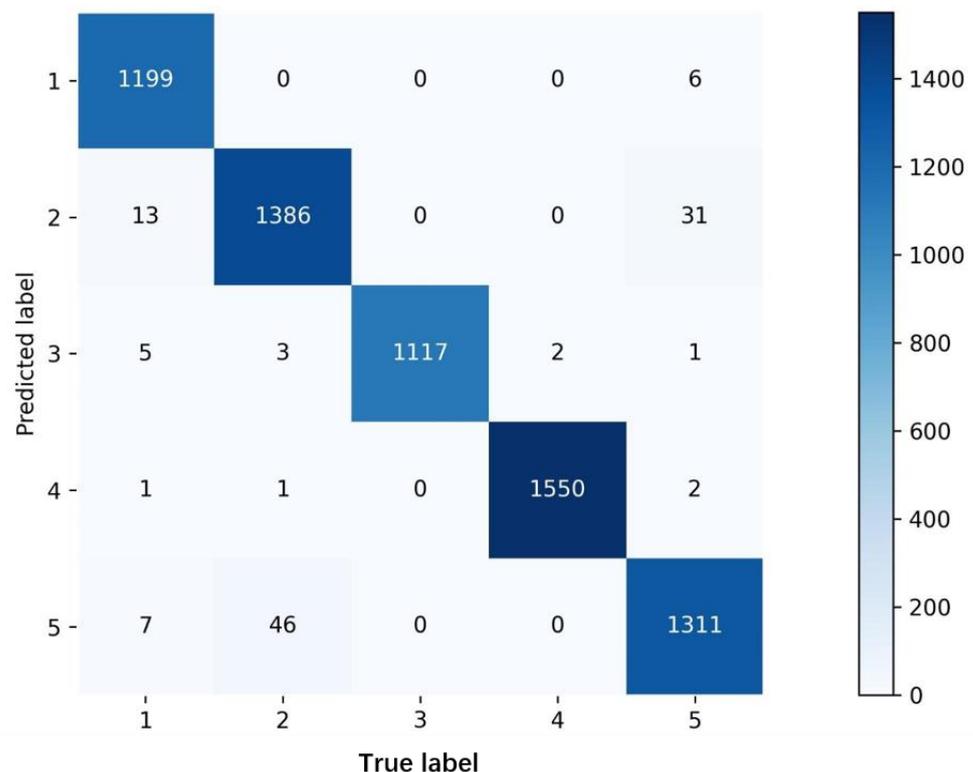


Figure 8. Identify the result confusion matrix. Note: label 1 Common rust; label 2 Gray leaf spot; label 3 Dwarf mosaic virus; label 4 Healthy; label 5 Northern leaf blight.

Taking the identification results of CLDD as an example, the misclassification samples of five types of leaves were sorted out. The visualization confusion matrix of misclassification of CLDD is shown in Figure 9, and the image position in the matrix is the misclassification samples. In the CLDD, the identification accuracy of label 3 reached 100% at the highest, and that of label 2 was 96.52% at the lowest, among which the misclassification rate of label 2 into label 5 was 3.20%. Combined with the test results in Figure 8, it was concluded that subjective vision in the process of corn leaf disease recognition that the disease texture and external morphological features were an essential basis for classification. When the disease texture features were relatively simple, and its texture features were distinguishable from other disease texture features to be identified, the classification and recognition accuracy of this disease was higher. When the difference in the texture features of the disease is slight, it is easy to misclassify. In Figure 9, the location samples corresponding to the visual confusion matrix of label 2 and label 5 showed that the phenotypic disease characteristics of label 2 showed irregular gray to long brown spots on the leaves, most of the leaves turned yellow and scorched, and label 5 had relatively little difference in the disease texture and external morphological characteristics. Hence, there were more misclassified samples in these two types of diseases. The test results in Figure 8

verified this conclusion. Further comparison and analysis of the test results in Table 4 and Figure 8 showed that the average recognition accuracy of healthy leaves in the CLDD was the highest. The reason may be that the sample pictures of the corn leaf dataset were single, primarily leaves, and the background was mostly simple pure color background, so the color and edge of the scanned leaves were clear.

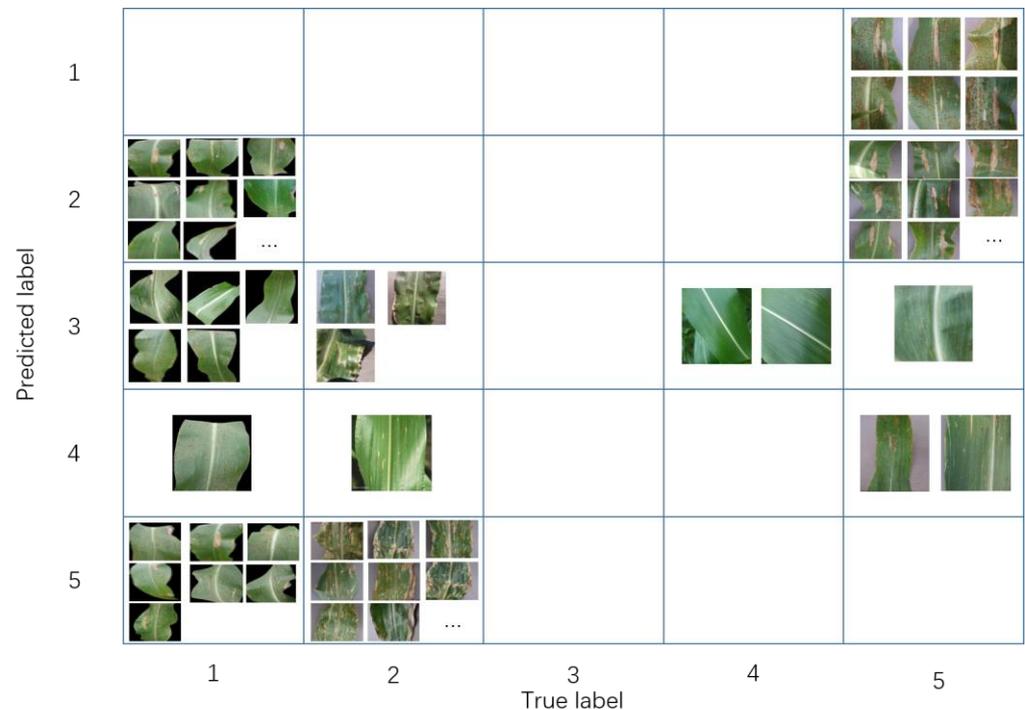


Figure 9. Misdivision visual confusion matrix. Note: label 1 Common rust; label 2 Gray leaf spot; label 3 Dwarf mosaic virus; label 4 Healthy; label 5 Northern leaf blight.

Table 4. Each kind of leaf accuracy on the CLDD.

Model	1	2	3	4	5	Accuracy (%)
Mobilenetv3	95.05	93.83	97.50	96.97	94.35	95.54
CD-Mobilenetv3	97.74	96.52	100.00	99.87	97.04	98.23

4. Discussion

The model proposed in our study compared with other DL models to identify plant diseases. Table 5 shows the performance comparison between the proposed model and other DL models. From Table 5, our proposed model had higher performance than most DL models.

Researchers usually collected datasets in the following ways: The first was through open-source datasets such as PlantVillage and Kaggle. The second was the online collection. The third was to collect and shoot in the field. Table 5 shows the accuracy of the CD-Mobilenetv3 model was higher than the improved lightweight models MobileNet-V2, MobileNet, Inception V3, and the heavyweight models VGG-16 and VGG. The accuracy of the improved AlexNet model proposed by Wang et al. was higher than the CD-Mobilenetv3 model. However, The CD-Mobilenetv3 model had fewer parameters, ran faster, and was more suitable for mobile terminal development. Based on the above data, the CD-Mobilenetv3 model had the best performance and was ideal for mobile devices.

Based on the Mobilenetv3 model, the CD-Mobilenetv3 model introduced the dilated convolution, Bias Loss function, and ECA module and added cross-layer connection between Mobile modules. The CD-Mobilenetv3 model not only improved the accuracy of

disease identification but also reduced the number of model parameters. There are some limitations in the CLDD, including only four kinds of diseased and healthy leaves. We will improve this aspect in future research.

Table 5. Performance comparison with other deep learning(DL) models.

Ref	Plant Species	Dataset	Acquired Method	Model	Accuracy (%)
Sun J. et al. [40]	4 kinds	10,371	Kaggle	Improved MobileNet-V2	92.20
Liu Y. et al. [44]	14 kinds	56,406	PlantVillage	MobileNet Inception V3	95.02 95.62
Xu J. et al. [26]	corn	5400	Online + Field collect	Improved VGG-16	95.33
Hassan, S.M. et al. [31]	3 kinds	20,020	PlantVillage + Field collect	Improved VGG	95.70
Wang, Y. et al. [32]	corn	10,785	Field collect	Improved AlexNet	99.35
Gao, Y. et al. [55]	wheat spikes wheat diseases	690 3754	Field collect	ResNet-50	85.56 99.32
Liu B Y. et al. [56]	Apples	5382	Field collect + Lab collect	MobileNet V2VGG	99.15 95.84
Zhang J L, et al. [57]	Lettuce	1918	Field collect	You Only Look Once v5	97.60
Proposed	corn	33,409	PlantVillage + ai-challenger + PlantifyDr + PlantDoc	CD-Mobilenetv3	98.23

5. Conclusions

Various leaf diseases occur in corn during planting, affecting corn's yield and quality. Our research proposed a model CD-Mobilenetv3 to identify corn leaf diseases efficiently and accurately. We experimented on the CLDD. The test results on the CLDD showed that the accuracy of the CD-Mobilenetv3 model proposed by our research reached 98.23%. The loss value reached 0.0285. The precision, recall, and F1 scores reached 98.26%, 98.26%, and 98.26%, respectively. The CD-Mobilenetv3 model compared several classic DL models ResNet50, ResNet101, ShuffleNet_x2, VGG16, SqueezeNet and InceptionNetv3, etc. On the CLDD, the accuracy increased by 1.62~6.39%. The loss value decreased by 0.0203~0.1854. The precision increased by 1.59~6.24%. The recall increased by 1.61~6.42%. The F1-score was 1.61~6.67% higher than other models. The parameter quantity was 1.12~134 M lower than most comparable models. All results verified the effectiveness of the CD-Mobilenetv3 model.

In future work, we will obtain more crop varieties and disease categories and study disease identification in the field environment. The group crops in the field environment are more challenging because of the complex environmental background. At the same time, we will try to apply the model to embedded devices for farmers to use.

Author Contributions: Conceptualization, C.B. and S.X.; methods, C.B. and S.X.; software, S.X.; validation, C.B. and S.X.; formal analysis, N.H.; survey, N.H. and S.Z.; resources, S.Z.; data management, Z.Z.; prepare the first draft, S.X.; writing comments and editors, C.B. and S.X.; visualization, S.X.; supervision, C.B.; project management, H.Y.; capital acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Development Program of Jilin Province (20220202032NC), China.

Data Availability Statement: Data are available from the author upon reasonable request.

Acknowledgments: The authors would like to thank the Jilin Agricultural University for their help with the experimental equipment provision for this study.

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

References

- Zhiyong, Z. Analysis and suggestions on the current situation of corn production in my country. *Grain Oil Feed. Sci. Technol.* **2022**, *3*, 4–8. (In Chinese)
- Manyi, D.; Qingkui, M.; Fang, G.; Biao, G.; Yaohua, H. Rapid identification of potato late blight based on machine vision. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 193–200. (In Chinese)
- Kai, X.; Qiliang, Y.; Chunxi, Y.; Xiaogang, L.; Huanhao, H.; Ping, Z. Predict Panax notoginseng disease incidence based on meteorological factors during a high disease incidence period. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 170–176. (In Chinese)
- Baofeng, S.; Yulin, L.; Yanchuan, H.; Rui, W.; Xiaofeng, C.; Dejun, H. Dynamic UAV remote sensing monitoring method for the incidence of population wheat stripe rust. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 127–135. (In Chinese)
- Peisen, Y.; Yifei, C.; Qianli, M.; Haoyun, W.; Huanliang, X. Research on the rice bacterial leaf spot identification method based on Random Forest. *Trans. Chin. Soc. Agric. Mach.* **2021**, *52*, 139–145+208. (In Chinese)
- Mahesh, B. Machine learning algorithms—A review. *Int. J. Sci. Res.* **2020**, *9*, 381–386.
- Shruthi, U.; Nagaveni, V.; Raghavendra, B.K. A review on machine learning classification techniques for plant disease detection. In Proceedings of the 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 15–16 March 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 281–284.
- Panigrahi, K.P.; Das, H.; Sahoo, A.K.; Moharana, S.C. Maize leaf disease detection and classification using machine learning algorithms. In *Progress in Computing, Analytics, and Networking*; Springer: Singapore, 2020; pp. 659–669.
- Shrivastava, V.K.; Pradhan, M.K. Rice plant disease classification using color features: A machine learning paradigm. *J. Plant Pathol.* **2021**, *103*, 17–26. [[CrossRef](#)]
- Xian, T.S.; Ngadiran, R. Plant diseases classification using machine learning. *J. Phys.* **2021**, *1962*, 012024. [[CrossRef](#)]
- Rong, Z.; Weiping, L.; Tong, M. A review of deep learning research. *Inf. Control.* **2018**, *47*, 385–397+410. (In Chinese)
- Liu, J.; Wang, X. Plant diseases and pests detection based on deep learning: A review. *Plant Methods* **2021**, *17*, 1–18. [[CrossRef](#)]
- Divakar, S.; Bhattacharjee, A.; Priyadarshini, R. Smote-DL: A deep learning based plant disease detection method. In Proceedings of the 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2–4 April 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6.
- Shruthi, U.; Nagaveni, V.; Arvind, C.S.; Sunil, G.L. Tomato plant disease classification using deep learning architectures: A review. In Proceedings of the 2nd International Conference on Advances in Computer Engineering and Communication Systems, Hyderabad, India, 11–12 August 2022; Springer: Singapore, 2022; pp. 153–169.
- Changji, W.; Qirui, W.; Hongrui, C.; Jianshuang, W.; Jun, N.; Ce, Y.; Hengqiang, S. Large-scale and multi-category pest identification model. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 169–177. (In Chinese)
- Ji, H.; Yu, J.; Lao, F.; Zhuang, Y.; Wen, Y.; Teng, G. Automatic position detection and posture recognition of grouped pigs based on deep learning. *Agriculture* **2022**, *12*, 1314. [[CrossRef](#)]
- Zhu, X.; Zhang, X.; Sun, Z.; Zheng, Y.; Su, S.; Chen, F. Identification of oil tea (*Camellia oleifera* C. Abel) cultivars using efficientnet-B4 CNN model with attention mechanism. *Forests* **2022**, *13*, 1. [[CrossRef](#)]
- Alokasi, H.; Ahmad, M.B. Deep learning-based frameworks for semantic segmentation of road scenes. *Electronics* **2022**, *11*, 1884. [[CrossRef](#)]
- Bangyal, W.H.; Rehman, N.U.; Nawaz, A.; Nisar, K.; Ibrahim AA, A.; Shakir, R.; Rawat, D.B. Constructing domain ontology for alzheimer disease using deep learning based approach. *Electronics* **2022**, *11*, 1890. [[CrossRef](#)]
- Pereira, T.D.; Tabris, N.; Matsliah, A.; Turner, D.M.; Li, J.; Ravindranath, S.; Papadoyannis, S.; Nirmad, E.; Deutsch, D.S.; Wang, Z.Y.; et al. SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **2022**, *19*, 486–495. [[CrossRef](#)]
- Bingzhen, L.; Ke, L.; Jiaojiao, G.; Wenzhi, J. A review of convolutional neural network research. *Comput. Age* **2021**, *4*, 8–12+17.
- Sun, H.; Zhang, S.; Ren, R.; Su, L. Surface defect detection of “Yuluxiang” pear using convolutional neural network with class-balance loss. *Agronomy* **2022**, *12*, 2076. [[CrossRef](#)]
- Zhang, B.; Zeng, K. A lightweight method for vehicle classification based on improved binarized convolutional neural network. *Electronics* **2022**, *11*, 1852. [[CrossRef](#)]
- Rao, D.S.; Ch, R.B.; Kiran, V.S.; Rajasekhar, N.; Srinivas, K.; Akshay, P.S.; Mohan, S.G.; Bharadwaj, B.L. Plant disease classification using deep bilinear cnn. *Intell. Autom. Soft Comput.* **2022**, *31*, 161–176.
- Koklu, M.; Unlarsen, M.F.; Ozkan, I.A.; Aslan, M.F.; Sabanci, K. A CNN-SVM study based on selected in-depth features for grapevine leaves classification. *Measurement* **2022**, *188*, 110425. [[CrossRef](#)]
- Jinghui, X.; Mingye, S.; Yichen, W.; Wenting, H. Image recognition of corn diseases based on convolutional neural network based on transfer learning. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 230–236+253. (In Chinese)
- Shougang, R.; Fuwei, J.; Xingjian, G.; Peisen, Y.; Wei, X.; Huanliang, X. Deconvolution guided tomato leaf disease identification and lesion segmentation model. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 186–195. (In Chinese)

28. Anagnostis, A.; Asiminari, G.; Papageorgiou, E.; Bochtis, D. A convolutional neural networks-based method for anthracnose infected walnut tree leaves identification. *Appl. Sci.* **2020**, *10*, 469. [[CrossRef](#)]
29. Maeda-Gutiérrez, V.; Galvan-Tejada, C.E.; Zanella-Calzada, L.A.; Celaya-Padilla, J.M.; Galván-Tejada, J.I.; Gamboa-Rosales, H.; Luna-García, H.; Magallanes-Quintar, R.; Guerrero Méndez, C.A.; Olvera-Olvera, C.A. Comparison of convolutional neural network architectures for classification of tomato plant diseases. *Appl. Sci.* **2020**, *10*, 1245. [[CrossRef](#)]
30. Wenxia, B.; Xuefeng, H.; Gensheng, H.; Dong, L. Recognition of corn leaf diseases based on improved convolutional neural network model. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 160–167. (In Chinese)
31. Hassan, S.M.; Jasinski, M.; Leonowicz, Z.; Jasinska, E.; Maji, A.K. Plant disease identification using shallow convolutional neural network. *Agronomy* **2021**, *11*, 2388. [[CrossRef](#)]
32. Wang, Y.; Tao, J.; Gao, H. Corn disease recognition based on attention mechanism network. *Axioms* **2022**, *11*, 480. [[CrossRef](#)]
33. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreeto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.
34. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, CA, USA, 18–23 June 2018; pp. 4510–4520.
35. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, CA, USA, 18–23 June 2018; pp. 6848–6856.
36. Türkmen, S.; Heikkilä, J. An efficient solution for semantic segmentation: Shufflenet v2 with dilated separable convolutions. In Proceedings of the Scandinavian Conference on Image Analysis, Norrköping, Sweden, 11–13 June 2019; Springer: Cham, Switzerland, 2019; pp. 41–53.
37. 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](#)]
38. Chao, X.; Sun, G.; Zhao, H.; Li, M.; He, D. Identifying apple tree leaf diseases based on deep learning models. *Symmetry* **2020**, *12*, 1065. [[CrossRef](#)]
39. Yang, L.; Guoqin, G. Using the improved SqueezeNet model to identify multiple leaf diseases. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 187–195. (In Chinese)
40. Jun, S.; Weidong, Z.; Yuanqiu, L.; Jifeng, S.; Yide, C.; Xin, Z. Identification of field crop leaf diseases based on improved MobileNet-V2. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 161–169. (In Chinese)
41. Shuqin, L.; Cong, C.; Tong, Z.; Bin, L. Plant leaf disease identification based on a lightweight residual network. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 243–250. (In Chinese)
42. 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](#)]
43. Eunice, J.; Popescu, D.E.; Chowdary, M.K.; Hemanth, J. Deep learning-based leaf disease detection in crops using images for agricultural applications. *Agronomy* **2022**, *12*, 2395. [[CrossRef](#)]
44. Yang, L.; Quan, F.; Shuzhi, W. Plant disease identification method and mobile application based on lightweight CNN. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 194–204. (In Chinese)
45. Xiaodong, Y.; Mengji, Y.; Haiqing, Z.; Dan, L.; Yiqian, T.; Xi, Y. Research and application of crop pest detection method based on transfer learning. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 252–258. (In Chinese)
46. Xiangpeng, F.; Yan, X.; Jianping, Z.; Zhilei, L.; Xuan, P.; Xiaorong, W. Grape leaf disease detection system based on transfer learning and improved CNN. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 151–159. (In Chinese)
47. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for Mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision on Seoul, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
48. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, CA, USA, 18–23 June 2018; pp. 7132–7141.
49. Hughes, D.; Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv* **2015**, arXiv:1511.08060.
50. Singh, D.; Jain, N.; Jain, P.; Kayal, P.; Kumawat, S.; Batra, N. PlantDoc: A dataset for visual plant disease detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD on Goa, Hyderabad, India, 5–7 January 2020; pp. 249–253.
51. Wei, Y.; Xiao, H.; Shi, H.; Jie, Z.; Feng, J.; Huang, T.S. Revisiting dilated convolution: A simple approach for weakly and semi-supervised semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, CA, USA, 18–23 June 2018; pp. 7268–7277.
52. Abrahamyan, L.; Ziatchin, V.; Chen, Y.; Deligiannis, N. Bias loss for mobile neural networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 6556–6566.
53. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. Supplementary material for 'ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 13–19.

54. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
55. Gao, Y.; Wang, H.; Li, M.; Su, W.-H. Automatic tandem dual blendmask networks for severity assessment of wheat fusarium head blight. *Agriculture* **2022**, *12*, 1493. [[CrossRef](#)]
56. Liu, B.Y.; Fan, K.J.; Su, W.H.; Peng, Y. Two-stage convolutional neural networks for diagnosing the severity of alternaria leaf blotch disease of the apple tree. *Remote Sens.* **2022**, *14*, 2519. [[CrossRef](#)]
57. Zhang, J.L.; Su, W.H.; Zhang, H.Y.; Peng, Y. SE-YOLOv5x: An optimized model based on transfer learning and visual attention mechanism for identifying and localizing weeds and vegetables. *Agronomy* **2022**, *12*, 2061. [[CrossRef](#)]

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.