

Research Article

Research on Strawberry Disease Diagnosis Based on Improved Residual Network Recognition Model

Xu Wenchao  and **Yan Zhi**

College of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China

Correspondence should be addressed to Xu Wenchao; xuwenchao@tjcu.edu.cn

Received 27 November 2021; Accepted 25 January 2022; Published 11 February 2022

Academic Editor: Hao Gao

Copyright © 2022 Xu Wenchao and Yan Zhi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Considering the problems of high cost, inefficiency, and time consumption of manual diagnosis of strawberry diseases, G-ResNet50 is proposed based on transfer learning and deep residual network for strawberry disease identification and classification. The G-ResNet50 is based on the ResNet50, and the focal loss function is introduced in G-ResNet50 to make the model devote itself to disease images that are difficult to classify. During the training process of the G-ResNet50 model, its convolutional layer and pooling layer inherit the pre-trained weight parameters from the ResNet50 model on the PlantVillage dataset, while adding dropout regularization and batch regularization methods to optimize the network model. The strawberry disease dataset includes four sample images of healthy plants, powdery mildew, strawberry anthracnose, and leaf spot disease. The dataset is enhanced and expanded by operations including angle rotation, adjusting contrast and brightness, and adding Gaussian noise. Compared with existing models such as VGG16, ResNet50, InceptionV3, and MobileNetV2, the results of model training and testing on 7,525 four-category leaf datasets show that the G-ResNet50 model has faster convergence speed and better classification effect, and its average recognition accuracy rate reached 98.67%, which is significantly higher than other models. Through the three evaluation indicators of precision rate, recall rate, and confusion matrix, it is concluded that the G-ResNet50 has good robustness, high stability, and high recognition accuracy and can provide a feasible solution for strawberry disease detection in practical applications.

1. Introduction

Disease problems are chiefly responsible for the decline of vegetable quality, which leads to farmers' economic losses, and are closely related to daily economic activities. As one of the main crops cultivated in greenhouse, strawberry also has many disease problems. How to quickly and accurately discover and identify strawberry diseases and take corresponding control measures is an important means to ensure strawberry growth, cure the disease, and increase farmers' income. Although different diseases show some symptoms in the visible light band, the information on their symptoms is complex and changeable, and only professionally trained plant experts can identify and diagnose these diseases. Common strawberry diseases include powdery mildew, strawberry anthracnose, and leaf spot, effectively solving the

problem of yield reduction caused by strawberry diseases, which mainly depend on accurately and quickly identifying strawberry diseases. Traditional manual identification of strawberry diseases is low in efficiency, poor in real time, has low accuracy, and is timeliness. Therefore, efficient and accurate identification of strawberry diseases can effectively reduce diseases and increase yield.

Only by discovering the primary symptoms of crop diseases as early as possible, the disease can be effectively eradicated and food security can be guaranteed. Numerous methods are proposed by lots of scholars and achieved remarkable results, and agriculture is gradually becoming a hot research field in the field of deep learning and agricultural disease recognition [1–3]. Deep learning and image recognition are effective in many research fields because they can provide more promising results at a lower cost.

With the great progress of deep learning and computer hardware, image recognition technology is applied to agricultural disease recognition by more and more scholars.

Various studies have shown that convolutional neural networks (CNNs) are extensively applied in detecting and classifying diseases. CNN does not use the previous manual feature extraction method, but automatically learns useful features from training dataset. CNN is proved the best-performing methods among many plant disease identification technologies [4]. Compared with previous machine learning methods with simple features, recent research is moving toward more accurate classifiers [5, 6]. Some researchers constructed a network based on CNN, and they proposed a method for identifying leaf diseases and extracting feature maps. The network model can diagnose and classify several diseases such as soybean leaf wilt, gray spot, and mosaic. In addition, this method uses feature maps to classify leaf diseases according to the severity of soybean leaves [7]. Literature [8] proposed a method for automatic selection and recognition network models based on CNN and transfer learning for many diseases of tomatoes. This method can match the neural network model that recognizes a certain disease with the highest similarity, which improved the accuracy of classification.

A network model for identifying peach leaf fungal punch hole disease, peach leaf shrinkage disease, and other diseases is proposed [9], which can automatically segment the purple-brown, dark-brown round, or irregular-shaped lesions in peach leaves, which have high recognition accuracy, especially for leaf corrugated bumps, gray-white powder on the surface of diseased leaves, etc. A large amount of experimental data shows that the classification rate is greatly improved compared with the comparison network models such as VGG and MobileV2, and the average classification accuracy rate reaches 97.75%. Rangarajan et al. in literature [10], respectively, used AlexNet and VGG16 for classification. The image dataset included six kinds of diseases and health images of tomato leaves. The classification accuracy of VGG16 is 99.24%, and the classification accuracy of AlexNet is 96.51%. In short, deep learning applications provide many opportunities in crop disease identification [10, 11].

In the literature [12–14], some researchers have proposed some methods that combine transfer learning to identify and classify diseases from plant leaf images. Sravan et al. in the study [12] discussed that the classic residual network model ResNet50 is used for training and verification on the public PlantVillage dataset, and the weight parameters are fine-tuned for several diseases' recognition, and the dataset contains more than 50,000 images of various plants, and the accuracy rate is as high as 99.26%. Study [13] used InceptionV3 model combined with the transfer learning method to carry out some research on apple leaf disease identification, and the average accuracy rate reached 97% on the PlantVillage dataset. Shin et al. [14] used six various CNN models to classify powdery mildew on strawberry leaves, such as VGG16, ResNet50, and other networks. Literature [15] applied the classic VGG16 network model to detect the diseases of millet crops. The studied disease images include two parts: training and testing. The

training dataset includes more than 1200 images, the test set includes nearly 200 leaf images, and the leaf images include disease and health categories. Finally, after fine-tuning the VGG16 network model and combining the parameter migration method, an accuracy of 95% is achieved. Literature [16] applied the transfer learning method to the trained convolutional neural network, fine-tuned it, and successfully realized the classification and recognition of medical images in lung, heart, gastrointestinal, and other diseases using the network model. The experimental results of above literature show that the method is superior to training a CNN from scratch in terms of classification accuracy and robustness.

PlantVillage is established by David Hughes, an epidemiologist at the University of Pennsylvania, and is currently the largest public plant dataset. The dataset includes more than 53,000 plant images of healthy plants and diseased leaves in 38 categories. Many network structure designers have proposed lots of different and improved models on the PlantVillage dataset, and most of the experimental results have achieved high accuracy. Marcel Salathé, the cofounder of PlantVillage, is planning to use machine learning technology to allow applications to automatically recognize plant species. The user only needs to upload the diseased plants to the database, and the machine learning algorithm can know what is wrong.

Some researchers [17–19] have applied augmentation techniques to augment the dataset by performing subtle mathematical transformations on geometric shapes, such as rotation angle, horizontal and vertical flip, zoom, size cropping, translation, and adding pepper noise and salt and Gaussian noise. These methods all make some changes based on the same image. It can better improve the overfitting problem and improve the predictive ability of the network model, and satisfactory results are obtained by these techniques. The scholars in [20] expanded the original dataset using translation, rotation, mirroring, and noise addition and applied the GoogLeNet model on the augmented dataset for training. After multiple epochs of experiments, the scholars completed the recognition and classification of 79 crop diseases among 14 different crop leaves on the validation set. Some preprocessing techniques are applied to transform the original dataset such as rotating, mirror, and Gaussian noise addition, so that the CNN model can more accurately and effectively identify and classify the original dataset [21]. Therefore, the construction of a broader representative dataset can solve problems such as model overfitting, and it is also one of the most important problems that the CNN method needs to solve in the identification of plant disease, and some more effective methods are needed to solve problems such as network model performance.

There are several other popular deep CNN models, such as AlexNet, GoogLeNet, VGG, InceptionV3, ResNet, and Xception. Deep CNN models still have some difficulties after the depth increases, such as the disappearance of gradients [22]. To solve these problems, residual modules can alleviate some of the problems. In the method proposed in this document [23], in the training process of the network, a momentum-driven stochastic gradient descent algorithm is selected as the optimizer of the back propagation, and good

results are achieved. In the study [24], the excellent ResNet model is used to detect tomato diseases, and several different improved ResNet structures are compared and tested, and it confirmed that ResNet50 is better than ResNeXt50. In literature [25], the ResNet model is combined with a faster R-CNN detector, and many comparative tests illustrate that the modified network has excellent performance in disease classification and identification of greenhouse crop leaves. Literature [26] used the ResNet50 model to identify grape yellow disease and compared other six CNN models (e.g., VGGNet and ResNet18). After the performance comparison test, the accuracy of classification rate of the ResNet50 reached 99.18%.

The above models all have the following problems as the number of model network layers increases, such as slowing down the convergence speed of the network and decreasing the accuracy of image classification [27–30].

In reference [31], the author proposed an improved deep residual 3D convolutional neural network model in response to the problems encountered in the automatic detection of rotten blueberries in the food industry. The blueberry spectral imaging mode is used to extract spectral and spatial features, the Parzen estimator is adaptively combined (TPE) and a variety of methods are chosen to optimize the network model, and the rapid identification and classification of blueberry sample images and the optimization of the parameters of the network model are realized. The classification accuracy of the network model reached 96.69%, and the convergence ability is also improved. In reference [32], the author proposed an improved ResNet model for the image classification of multiple diseases in medical X-ray images, the global average pooling is replaced by an adaptive dropout of medical image classification, and the multi-label classification is converted into N two meta-classification, and the last multiple experimental evaluations show that compared with the traditional architecture and VGG16, the accuracy of the proposed model achieved 87.71% and 81.8% on different datasets, respectively. The author in the literature [33] describes a large number of algorithms and applications related to data analysis, pattern recognition, machine learning, deep learning, and the Internet of things, especially in the field of health care; for example, various advanced neural network algorithms are used to identify and classify disease images.

Some research literature also shows that residual network (ResNet) can effectively improve the problem of reduced training effect as the number of convolutional layers increases, speed up the convergence, and achieve better recognition results compared with other networks [34–36]. At present, the ResNet can replace the VGG network as the basic feature extraction network in general computer vision problems. The trained network after transfer learning can be applied to similar fields, to obtain better feature extraction capabilities and improve the classification capabilities.

An improved residual network G-ResNet50 is used to identify healthy strawberry plants, powdery mildew, strawberry anthracnose, and leaf spot disease images in this study. The focal loss function is introduced in the G-ResNet50 to make the model devote itself to disease images that are difficult to classify, the weight parameters learned on the PlantVillage dataset are transferred to the G-ResNet50, and finally dropout

regularization and batch regularization methods are adopted to optimize the G-ResNet50 model, and the construction of the strawberry disease image recognition model is completed.

2. Materials and Methods

2.1. Materials and Equipment. The selected dataset comes from the strawberry disease images used in the strawberry planting base of Tianjin Academy of Agricultural Sciences. The strawberry disease image dataset contains sample images of healthy strawberry, powdery mildew, strawberry anthracnose, and strawberry leaf spot disease. After data enhancement and expansion, there are a total of 7525 RGB disease image samples. Among them, there are 1140 images of healthy strawberry, 2230 images of powdery mildew disease, 1975 images of strawberry anthracnose disease, and 2180 images of strawberry leaf spot disease. The samples of the strawberry disease image are shown in Figure 1.

Firstly, the preprocessing operation is performed on the strawberry disease images in the dataset, and the repeated and useless disease images are eliminated. The preprocessed strawberry disease image is divided into three parts: training, verification, and testing, according to the sample label, and it is divided into a ratio of 6 : 2 : 2. The strawberry disease image samples are labeled, and the label array adopts 2D one-hot encoding labels.

Secondly, considering that preventing the overfitting problem and the disease image dataset is expanded by operations such as flipping, rotating, and adding noise, the following operations are performed on the disease images: (1) image brightness and contrast adjustment: the image is randomly adjusted 0.7–1.1 times the original brightness and contrast. (2) Image horizontal and vertical flip: the image is randomly flipped in horizontal or vertical direction. (3) Random rotation, transformation, or reverse transformation. (4) Zoom: the random zoom ratio is 0.9–1. (5) Artificial noise: Gaussian noise is added. Data augmentation is shown in Figure 2.

Finally, all diseased images are adjusted to 224×224 pixels, and then, batch normalization (BN) processing is performed on the image as the input layer. Batch normalization can not only increase the convergence speed and improve the ability of the network to generalize better, but also effectively alleviate the problem of gradient disappearance.

This research used 64 bit Microsoft Win10 OS, Intel Core i9-9900KS CPU @ 3.20 GHz, 64 GB RAM, 2T hard drive, and 8G NVIDIA GeForce GTX 2080 Ti GPU. Python version is 3.7. Keras version is 2.3.1.

2.2. Deep Residual Network. ResNet50 is a residual learning framework that reduces the burden of network training. This kind of network is inherently deeper than the previously used network. Residual learning is to add a shortcut connection on the basis of the traditional linear convolutional network structure, thus enhancing the feature transfer. In the training process, the error of the bottom layer can be passed to the next layer through the shortcut connection,

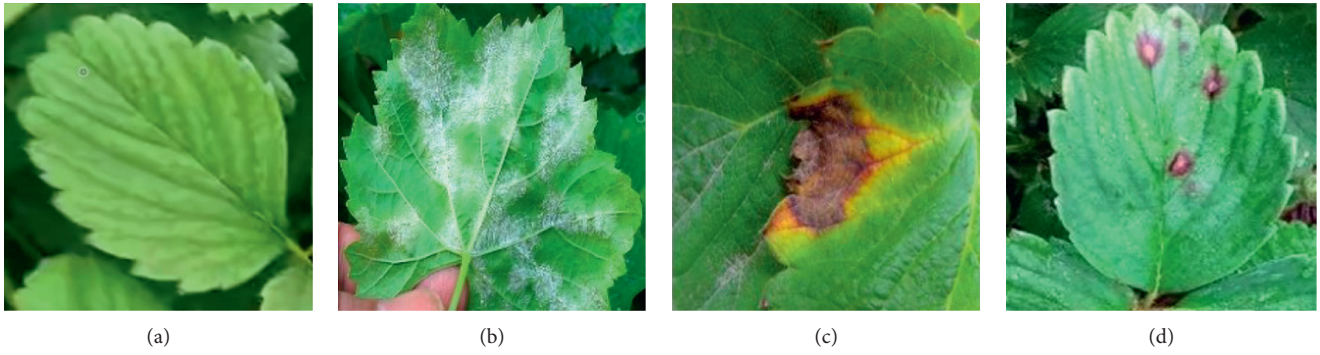


FIGURE 1: Samples of the strawberry disease image. (a) Healthy, (b) powdery mildew, (c) strawberry anthracnose, and (d) leaf spot.

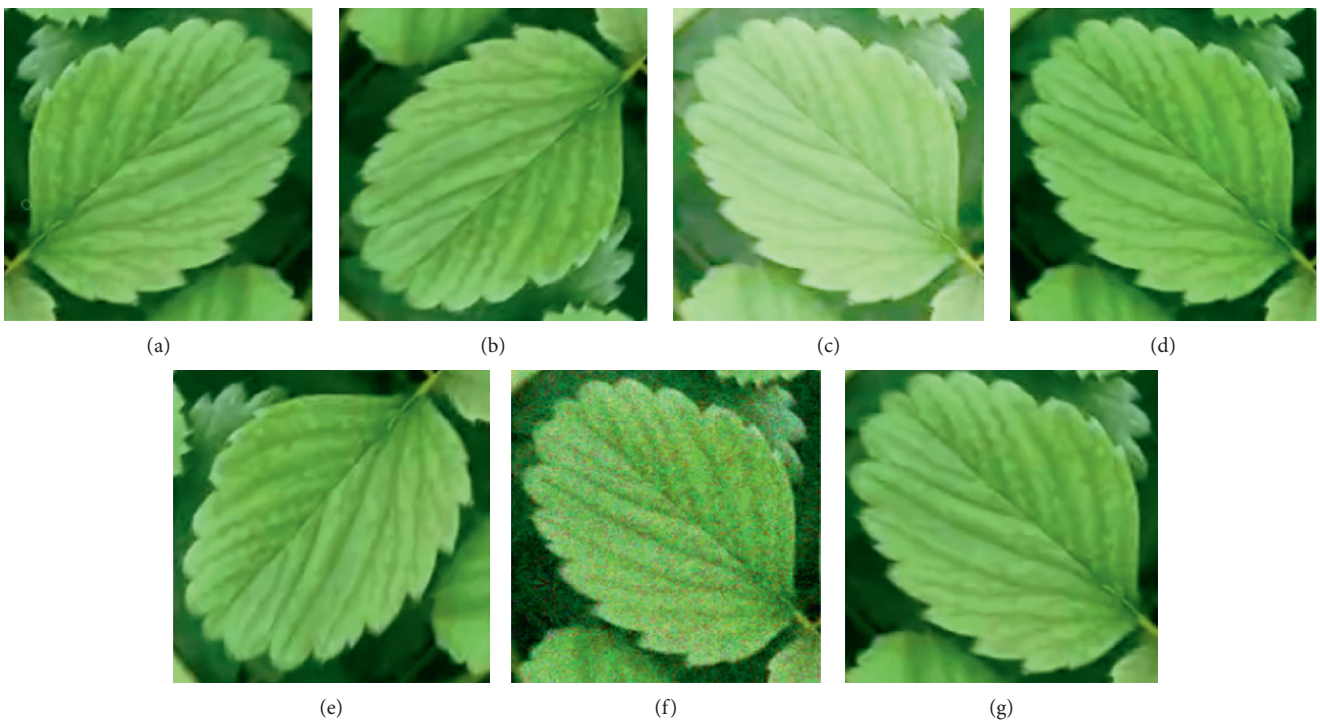


FIGURE 2: Data augmentation. (a) Flip vertically; (b) flip horizontally; (c) adjust brightness; (d) adjust contrast; (e) rotate 90 degrees counterclockwise; (f) adding Gaussian noise; and (g) zoom.

and the convergence speed of the network will be accelerated, and more layers of network training can be carried out to achieve higher classification accuracy. As the network layers increase, the accuracy of most network models will tend to be saturated and will decline rapidly. He Kaiming, the designer of ResNet, believes that this phenomenon is caused by overfitting and the increase in the number of model network layers, which lead to increased training errors. The core idea of ResNet is to calculate the residuals from layer to layer and train the deep network with a simpler and clearer framework.

A deep residual learning framework is introduced into the deep residual network model for solving the degradation problem. For deep-level models, additional layers are constructed through identity map, and let these layers fit a residual map.

Assuming that stacked layers fit another residual map $F(x) = H(x) - x$, the original map can be expressed as $F(x) - x$, and it is realized through adding shortcut connections, and it can implement identity mapping and add its output to the output of the stacked layer. The schematic diagram of residual learning is shown in Figure 3. The experiment proves that the introduced shortcut connection does not add additional parameters, nor does it affect the complexity of the original network.

2.3. Transfer Learning. Due to the small size of the strawberry leaf disease image dataset, using the dataset directly for the training of the network model will cause problems such as overfitting or low recognition accuracy. Transfer learning is to transplant the weight parameters trained on other large

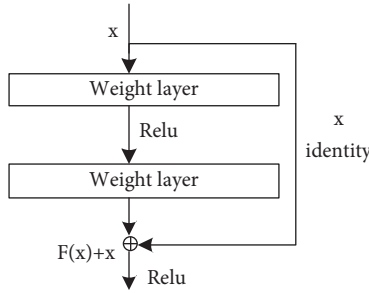


FIGURE 3: Schematic diagram of residual learning.

datasets to the network model to be trained, which can improve the training efficiency of the model. Parameter transfer is to randomly initialize a few convolutional layers of the network model, and weight parameters of the pre-trained network are transplanted to other convolutional layers, and the network parameters are retrained through the dataset to be trained.

In this proposed network, the parameter transfer method is applied to the training network, and the weight parameters trained on the large dataset PlantVillage are transferred to the improved residual network for the recognition of strawberry leaf diseases.

2.4. Loss Function. In view of the imbalance in the classification and the difference in mining difficult samples, the focal loss function is proposed. The object of this research is strawberry disease images, and the trained model also needs the ability to distinguish whether strawberry is healthy. Therefore, the training dataset also contains healthy strawberry image samples, and healthy samples are easier to classify than disease samples. The focal loss function adds a factor on the original function, the weight of easy-to-divide samples and the loss are reduced during network model training, and the difficult and misclassified samples are focused.

The focal loss function is modified based on cross-entropy loss function. After this function is regressed by the softmax function in the deep residual network, its expression is as follows:

$$CE(x) = -\lg\left(\frac{e^{x_i}}{\sum_j e^{x_j}}\right), \quad (1)$$

where i and j represent the category number, and x is the characteristic value. The modified focal loss function expression is as follows:

$$FL(\mathbf{p}_t) = -(1 - \mathbf{p}_t)^\gamma \lg(\mathbf{p}_t). \quad (2)$$

In equation (2), γ represents the focus parameter, and $\gamma \geq 0$, $(1 - \mathbf{p}_t)^\gamma$ represents modulation factor, which is used to reduce the weight of images that are easier to identify. The strawberry disease image recognition network can focus on images that are more difficult to identify during training.

2.5. Model Establishment. Transfer learning and focal loss function are introduced to construct a strawberry disease image recognition model based on the ResNet50 structure, which is called G-ResNet50.

First, the ResNet50 is pretrained on the PlantVillage dataset, the pre-trained weight parameters are obtained and saved, and then the parameters are migrated to the G-ResNet50.

The PlantVillage dataset contains more than thirty categories of leaves according to types and diseases. The PlantVillage dataset contains many types of plants, such as grape, citrus, potato, soybean, strawberry, apple, blueberry, pumpkin, strawberry, peach, and tomato. It can greatly help G-ResNet50 model to learn the disease characteristic information of various plant leaves. A sample image of the PlantVillage dataset is shown in Figure 4.

Figure 5(a) shows the architecture diagram of ResNet50. Figure 5(b) shows the Res block structure. ResNet is divided into 5 stages. Among them, Stage 0 has a relatively simple structure and can be regarded as a preprocessing of input. The last four stages are all composed of Res blocks, and their structures are relatively similar. The following is a detailed description of Stage 0 and Stage 1, and the principles of the following three stages are similar. The input of Stage 0 is $(3 \times 224 \times 224)$, which, respectively, indicates the number of channels, height, and width of the input. The first layer in Stage 0 includes 3 sequential operations: (1) convolution, where the size of the convolution kernel is 7×7 , and the number of convolution kernels is 64; (2) batch normalization; and (3) ReLU activation function. The second layer in Stage 0 is the largest pooling layer, with a kernel size of 3×3 and a step size of 2. $(64, 56, 56)$ represents the number of channels, height, and width of the stage output, where 64 represents the number of convolution kernels in the first convolution layer of the stage, and 56 represents the shape size. Generally speaking, in Stage 0, the input of shape $(3 \times 224 \times 224)$ performs operations such as convolutional layer, BN layer, ReLU activation function, and Max Pooling and then obtains output of $(64, 56, 56)$. The input shape of Stage 1 is $(64, 56, 56)$, and the output shape is $(64, 56, 56)$.

The focal loss function is introduced to estimate the loss of the G-ResNet50, and Figure 5 is the final network transfer learning architecture diagram.

As can be seen from Figure 5(c), the strawberry disease image of the dataset is used as the input of the G-ResNet50, and it first passes through a 7×7 convolutional layer, then performs batch normalization and activation, and then passes through the 3×3 maximum pooling layer, 4 residual learning blocks, and an average pooling layer, and then, the multidimensional output is converted into one-dimensional through the flatten layer. The one-dimensional nonlinear combination features are learned through the fully connected layer, and two fully connected layers are set up, where fully connected layer (FC1) is 1000, and FC2 is 4. Due to the limited experimental samples, to avoid the occurrence of overfitting, a dropout layer is added after the fully connected

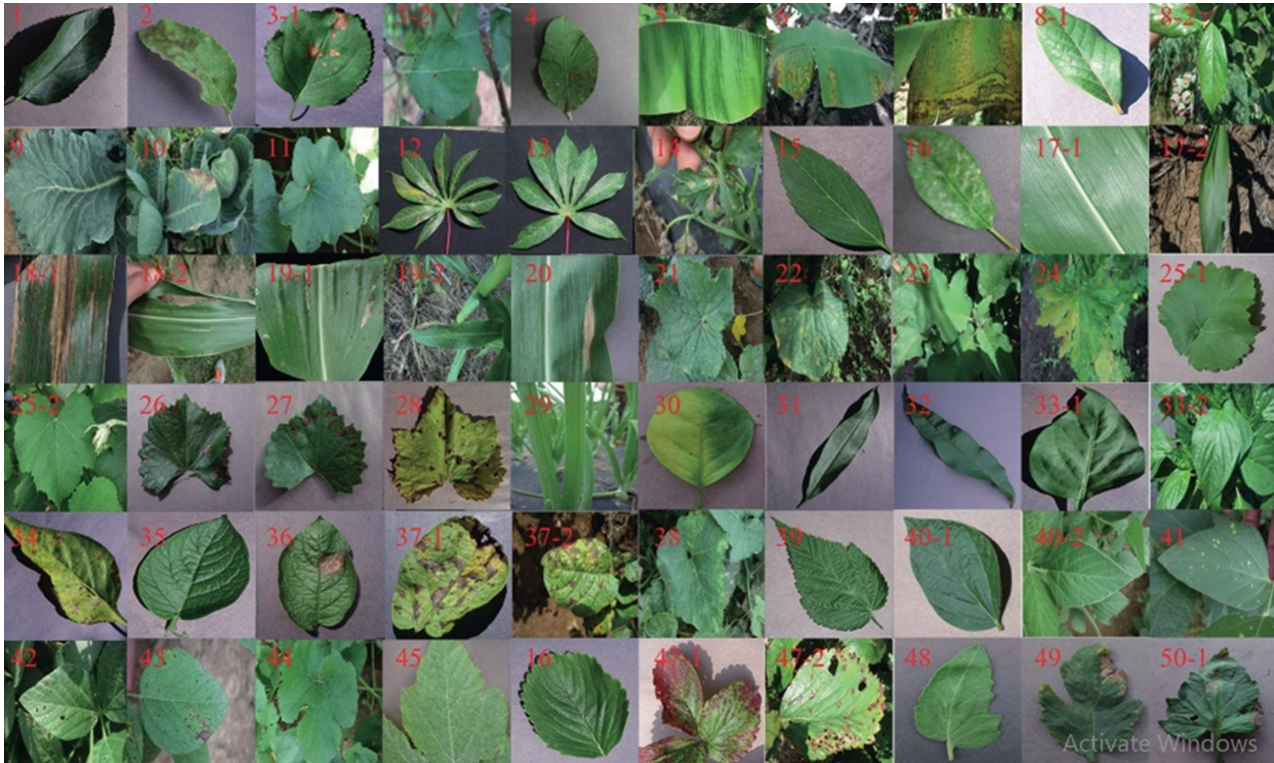


FIGURE 4: Sample image of the PlantVillage dataset.

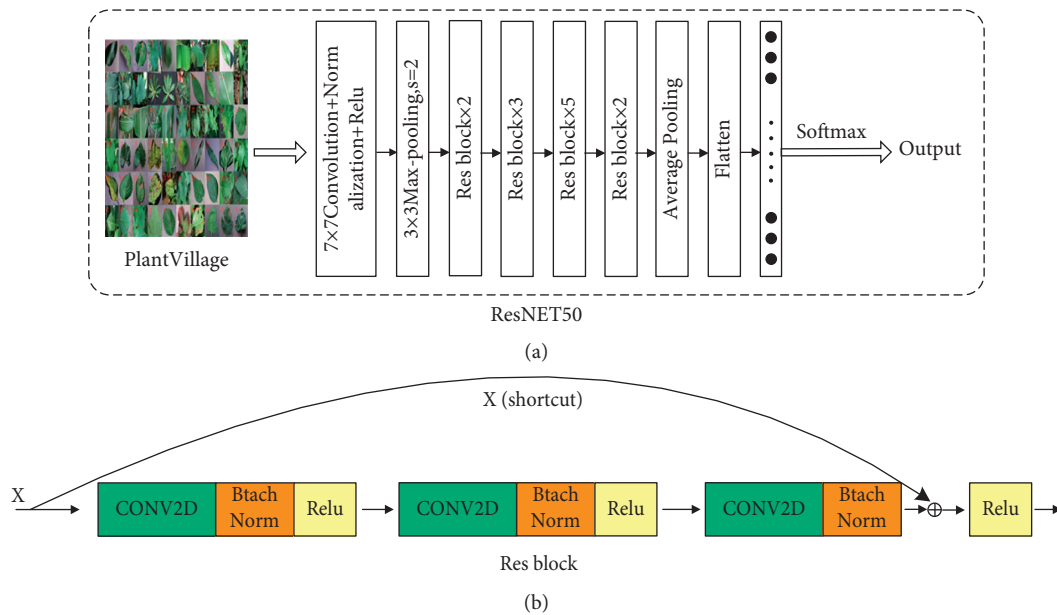


FIGURE 5: Continued.

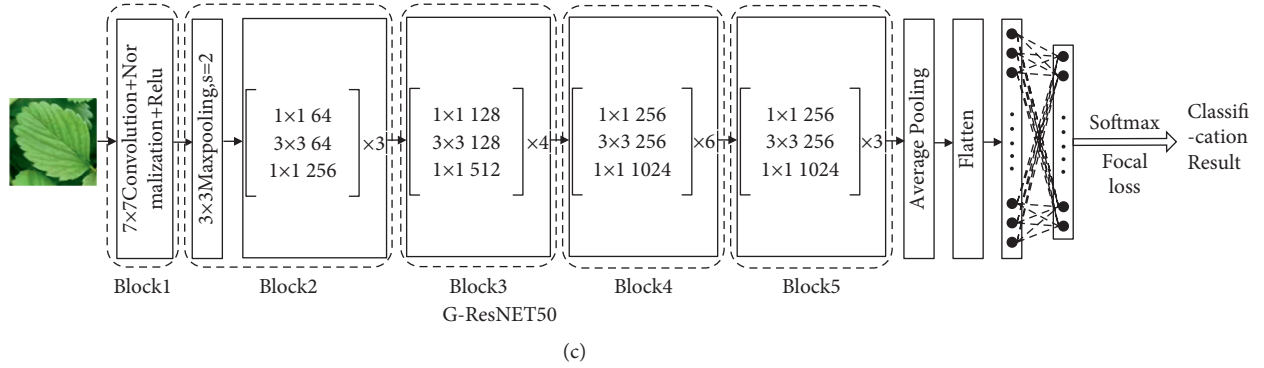


FIGURE 5: G-ResNet50 transfer learning architectural diagram. (a) Architectural diagram of ResNet50. (b) Res block architecture in ResNet50. (c) Architectural diagram of G-ResNet50.

layer, and dropout rate is set to 0.5; that is, half of the parameters are randomly discarded, and at the same time after each fully connected layer, the linear correction unit (ReLU) is introduced to solve the gradient dispersion problem and finally is classified by the softmax classifier and output prediction results.

The core of the G-ResNet50 is the residual learning block based on the ResNet50. These residual blocks are represented as follows:

$$y = F(x, W_i) + x. \quad (3)$$

y represents the output vectors of the current layer of the G-ResNet50, and the function $F(x, W_i)$ is the residual function relationship to be learned. When the dimensions of x and F are not equal, the residual block is represented as follows:

$$y = F(x, W_i) + W_s x. \quad (4)$$

W_s represents the linear projection performed by the shortcut connection, and W_s is only used when matching dimensions. In the G-ResNet50, Conv2 layer, Conv3 layer, Conv4 layer, and Conv5 layer contain 3, 4, 6, and 3 residual blocks, respectively. The residual block structure in the G-ResNet50 is shown in Figure 6 (the residual block in Figure 6 is the residual block for the first tensor dimension conversion, so there is a dimensional conversion convolutional layer on each shortcut connection). BN and ReLU are used in each residual block, so that after processing, the residual learning unit will be easier to train and more generalized than before.

3. Experimental Results and Discussion

3.1. Model Training. The G-ResNet50 is used to test 7525 RGB image samples and applied G-ResNet50 and other classification network models (ResNet50, VGG16, InceptionV3, and MobileNetV2) for comparison experiments. Each update of the gradient value adopts the SGD algorithm, in which the initial learning rate of SGD optimizer is set to 0.0001. Since each epoch of the stochastic gradient descent method will have relatively large noise and easily fall into the local optimal solution, momentum is used to optimize the

SGD optimizer. Momentum will keep the previous direction while updating the direction, which can increase the stability of the model and has the ability to get rid of the local optimum. The parameter momentum value is assigned to 0.9. After using Nesterov, the learning rate decay value is assigned after parameter is updated to 0.0001. The value of batch size for each processing is set to 32. The best-performing model is saved using the ModelCheckpoint method during training. Cross-entropy error function measures the difference in distribution in the process of learning the networks in comparative test. Finally, the parameter epoch value is set to 35 for training.

3.2. Model Feature Extraction and Visualization. Strawberry disease images are randomly selected and the feature extraction process of the convolutional layer of the G-ResNet50 is observed. The visualization of the convolution kernel is shown in Figure 7.

The visualization of the convolutional neural network model can help intuitively understand the classification model. Through the visualization operation, it is possible to observe the feature changes in the convolutional layer in the middle of the network structure, and it is possible to know what disease features have been learned by each layer of the network structure. Through the learned disease characteristics, the network parameters can be further adjusted to enhance the accuracy.

Figure 7 shows the extraction effect of the output feature maps of each layer after the model training is completed. It is clear that the learning of the convolutional neural network is hierarchical. It can be seen from the disease feature map of Conv1_x that the shallow layer is usually a collection of edge filters, presenting some mixed information such as edges and colors. Almost all the information in the original image is retained. From the output feature maps of Conv2_x and Conv3_x, it can be seen that as the depth of the convolutional layer increases, there is less and less information about the visual content, and more feature information such as the texture of the disease image is presented. At a deeper level, the feature map outputs by the Conv4_x and Conv5_x convolutional layers are becoming more and more abstract, and the final features all represent the relevant feature

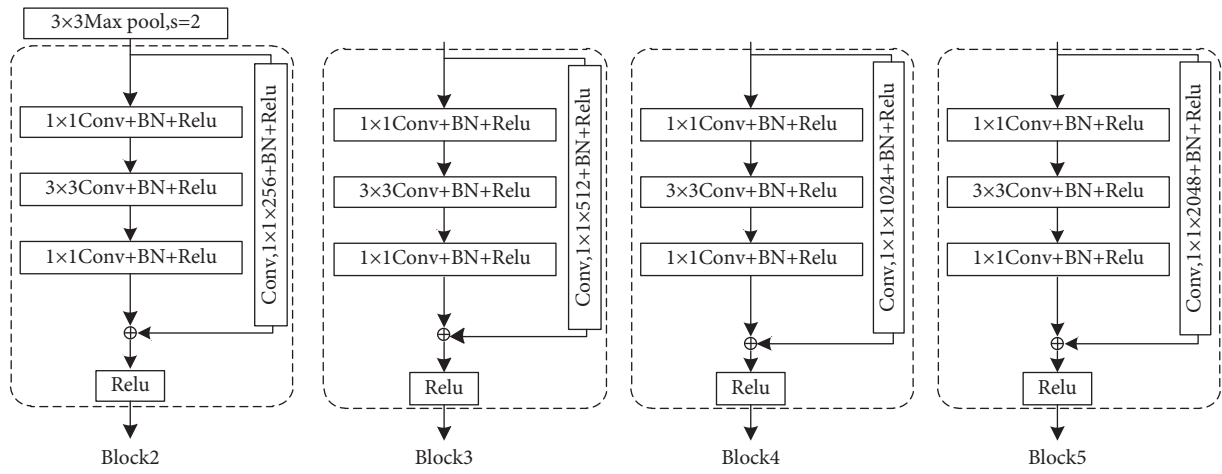


FIGURE 6: Residual block structure in G-ResNet50.

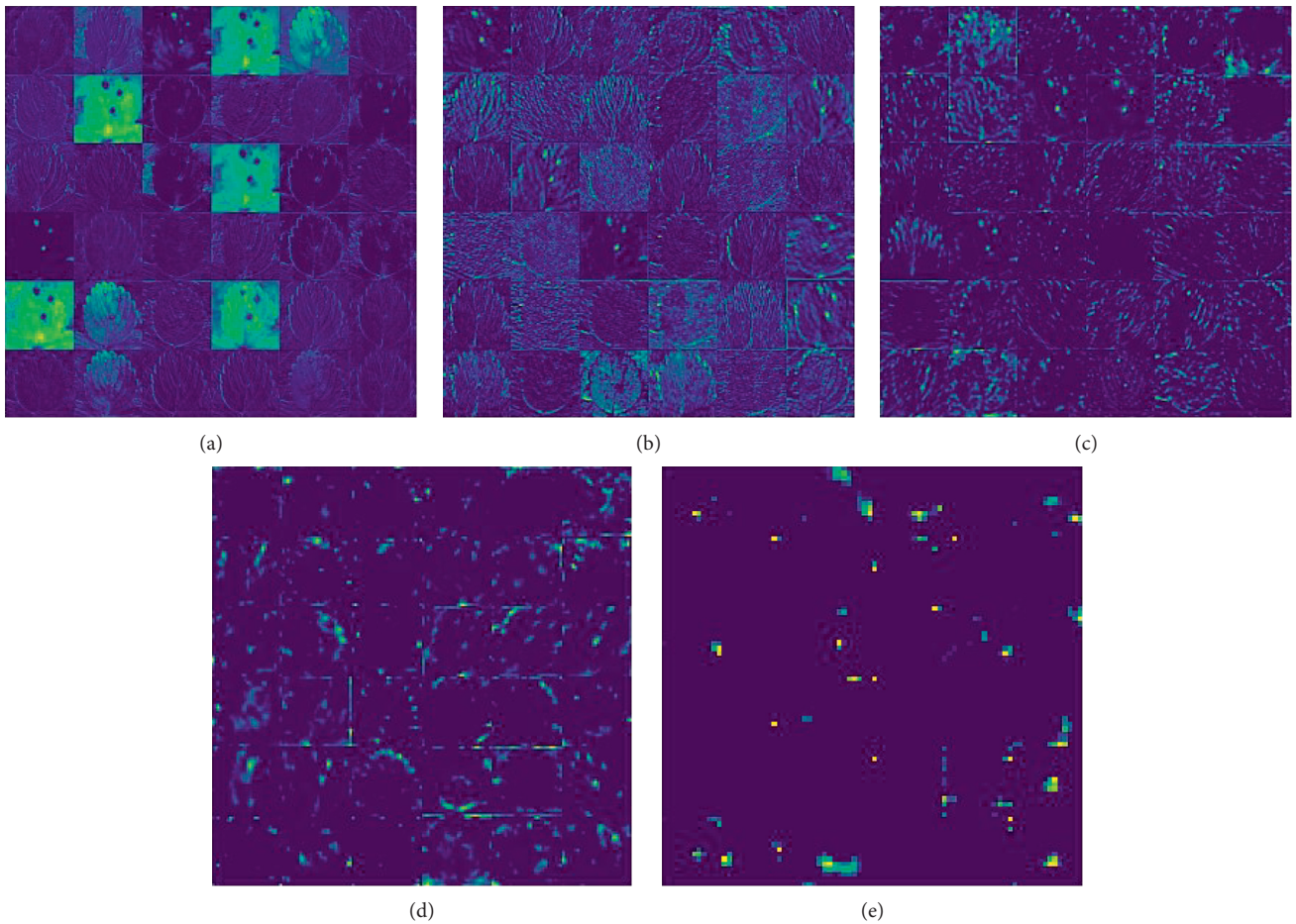


FIGURE 7: Convolution kernel visualization of strawberry disease image. (a) Conv1_x; (b) Conv2_x; (c) Conv3_x; (d) Conv4_x; and (e) Conv5_x.

information of different categories of disease, which are used to make the final classification of the model.

Through the visualization process of the feature map, it can be found that the visualization of the feature map is very useful in the model reproduction process and can be used to locate the error of the model. It can be summarized as follows:

- (1) The shallow network extracts texture and detailed features
- (2) The deep network extracts the strongest features such as contour and shape
- (3) The shallow network contains more features and also has the ability to extract key features (e.g., the eighth feature map in the Conv1_x feature map, which extracts strawberry disease features)
- (4) Relatively speaking, the deeper the number of layers, the more representative the extracted features, and the resolution of the image is getting smaller and smaller

It can be seen that the features extracted from different feature maps are almost different, and some focus on the edge, while others focus on the whole. Compared with the deeper features, the shallow features are mostly complete, while the deeper features are smaller and smaller. Of course, these feature maps are all related to each other, and the network structure is a whole.

According to the visualization of the output features of the middle layer, it can be seen that the improved model can well extract strawberry disease characteristics, which can be used to classify and recognize strawberry health, powdery mildew, anthracnose, and leaf spot disease images.

3.3. Experimental Results and Discussion. After setting the parameters, the strawberry disease image dataset is trained on five network models: G-ResNet50, ResNet50, VGG16, InceptionV3, and MobileNetV2, and each network is trained for 35 epochs. Figure 8 shows the training results of the five networks. According to the accuracy of the dataset, the recognition effect of strawberry disease images from high to low is as follows: G-ResNet50, ResNet50, InceptionV3, VGG16, and MobileNetV2, and each network model is already closed complete convergence in the 10th epoch. It can be found from Figure 8 that the most stable model on training set loss is the G-ResNet50, and the most unstable model is the MobileNetV2, and the loss of each model on the validation set fluctuates slightly. After 35 epochs' training, the loss value of each model reached a very low value, but the loss value of the MobileNetV2 on verification set is the most unstable, and the loss curve changes are the most. Finally, the trained model is tested, and the average recognition rate of each model on the test set is obtained. Among them, the G-ResNet50 reached 98.67%, ResNet50 reached 97.74%, InceptionV3 network reached 97.27%, VGG16 network reached 96.21%, and MobileNetV2 reached 95.81%.

Compared with the other four comparison models, the G-ResNet50 has a higher average recognition rate, faster

model convergence, and better robustness and generalization capabilities.

The performance of the test network is evaluated separately, and the precision rate (precision), the recall rate (recall), and F1 score are used as the evaluation indicators of G-ResNet50. First, the following settings are made, TP represents the number of real positive samples, FP represents the true number of negative samples, FN represents the number of false-negative samples, and the precision, recall, and F1 score are, respectively, are represented as follows:

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP}, \\ \text{recall} &= \frac{TP}{TP + FN}, \\ F1 - \text{score} &= 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \end{aligned} \quad (5)$$

After predicting the strawberry disease images on the test set, the evaluation indicators of each model for the four samples are shown in Table 1. These evaluation indicators include the precision, recall, and F1 score. 4 types of strawberry disease images dataset include healthy strawberry plants (hea), powdery mildew (p-m), strawberry anthracnose (ant), and leaf spot (l-s) images.

The precision represents the accuracy of the model's recognition of sample images. The G-ResNet50 recognition accuracy of the four sample images reached 99.56%, 98.63%, 99.24%, and 97.72%, respectively, which is the most effective of all network models. The higher the precision, recall, and F1 score, the better the effect of the model. As is known from the comparative test, the G-ResNet50 has better recognition effect and generalization ability.

A confusion matrix is also one of the evaluation indicators of the classification model. The confusion matrix of the visual test set on each model is shown in Figure 9.

The column label of the confusion matrix represents the predicted category, and the sum of the corresponding row values represents all number of images in the category. The row labels of the confusion matrix represent the true category of the predicted image.

The value at the intersection of the row and column represents that the category is predicted as the number of corresponding column labels. The value at the diagonal line shows correctly predicted tags; the darker the diagonal line indicates the better the model's effect. The confusion matrix of each model illustrates the recognition ability of the four types of strawberry disease images.

The lesions of different diseases have a certain similarity at different times and different angles, so there are cases of identification errors.

Through the confusion matrix of the five networks, it can be found that in the G-ResNet50, InceptionV3, ResNet50, MobileNetV2, and VGG16 network model, the number of leaf spots correctly identified as leaf spots is 429, 417, 419, 419, and 424, respectively, and the number of leaf spots incorrectly identified as powdery mildew is 9, 13, 10, 28, and 33, respectively. The G-ResNet50 model has the highest

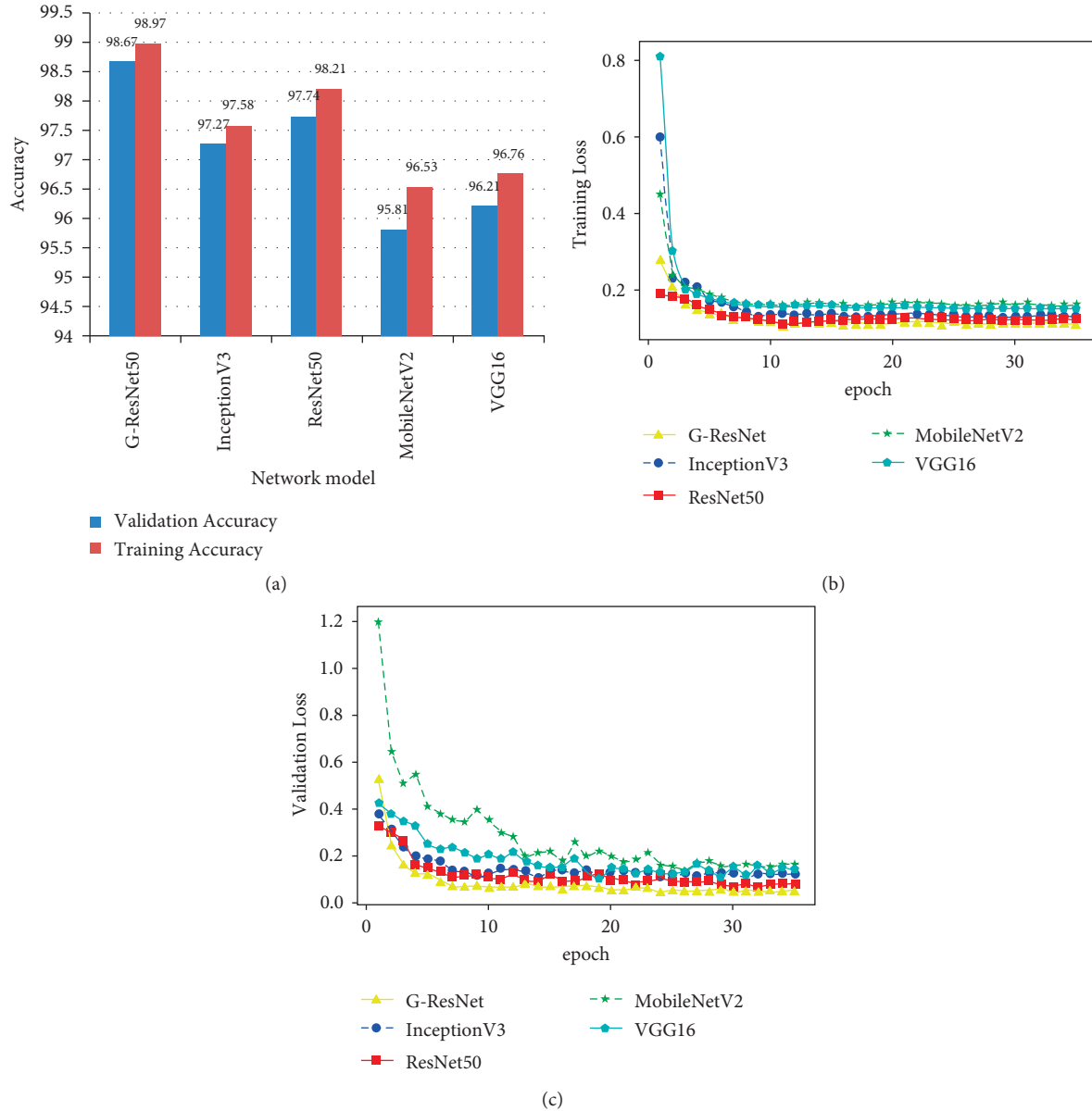


FIGURE 8: Accuracy and loss of training set and validation set during iteration of different models. (a) Accuracy of five network models. (b) Training loss of five network models. (c) Validation loss of five network models.

TABLE 1: Accuracy, recall, and F1 score of each model.

Model	Precision (%)				Recall (%)				F1 score			
	hea	p-m	ant	l-s	hea	p-m	ant	l-s	hea	p-m	ant	l-s
G-ResNet50	99.5	98.6	99.2	97.7	100	97.5	99.5	98.4	99.8	98.1	99.4	98.1
InceptionV3	99.6	95.8	98.4	96.5	98.7	96.6	99.0	95.6	99.1	96.2	98.7	96.1
ResNet50	99.5	96.4	98.7	97.2	99.1	97.3	99.2	96.1	99.3	96.9	99.0	96.6
MobileNetV2	97.3	96.0	97.7	93.1	97.8	92.2	98.5	96.1	97.6	94.1	98.1	94.6
VGG16	97.0	97.8	98.7	92.2	99.1	91.5	98.7	97.2	98.0	94.6	98.7	94.6

number of correct identifications, and the number of incorrectly identified leaf spots is the least, and the number of powdery mildew being misidentified is also the least.

The false recognition rate of these models is higher for strawberry powdery mildew and strawberry leaf spot disease.

Among them, the false recognition rate of VGG16 is the highest for strawberry leaf spot disease, and the false recognition rate of InceptionV3 is the lowest for strawberry powdery mildew disease. Compared with other models, the confusion matrix proves that the G-ResNet50 has the best

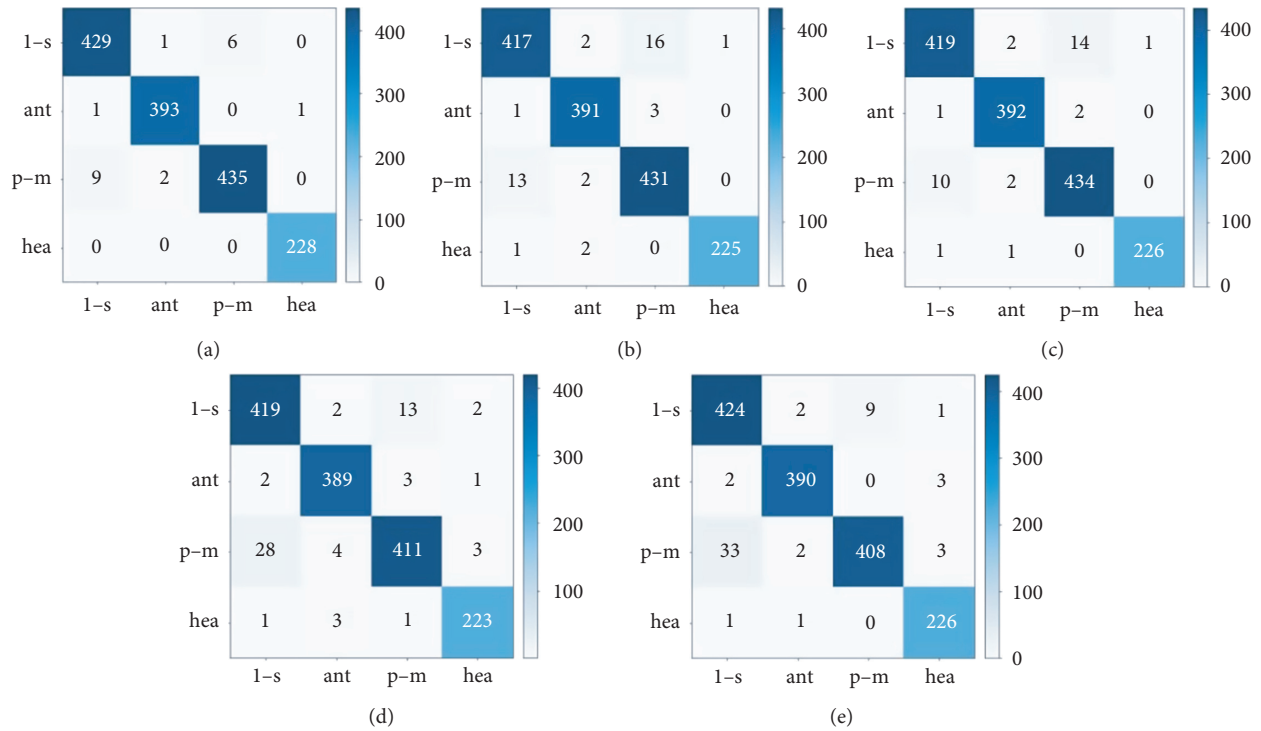


FIGURE 9: Confusion matrix for each model to identify strawberry diseases. (a) G-ResNet50; (b) InceptionV3; (c) ResNet50; (d) MobileNetV2; and (e) VGG16.

TABLE 2: Average classification accuracy of the five networks.

Network model	Dataset		
	Tianjin Academy of Agriculture Sciences	PlantVillage	IDADP
G-ResNet50	98.67	97.45	96.35
VGG16	96.21	95.31	94.46
ResNet50	97.74	96.78	95.73
InceptionV3	97.27	96.06	94.26
MobileNetV2	95.81	93.73	9.28

performance and the highest average recognition rate, which can provide a feasible solution for strawberry disease image recognition. In short, it is certainly hoped that our model can recognize the number of true positive and true negative that can appear more, and the number of false positive and false negative can be avoided as much as possible.

The PlantVillage dataset and the Image Dataset for Agricultural Diseases and Pests (IDADP) are two publicly available datasets about plant diseases, which can be used to evaluate the classification performance of different deep learning algorithms. The IDADP is constructed by the Hefei Institute of Intelligent Machinery, Chinese Academy of Sciences, and integrates a large number of agricultural pests and disease image sample resources, each of which has hundreds to thousands of high-quality images. The dataset has a complete range of pests and diseases, and the data volume exceeds 200 G. The PlantVillage dataset is a well-

known dataset about plant diseases, including many kinds of plant diseases and healthy leaf data.

The samples used in the experiment are all image-enhanced images. The number of strawberry disease images contained in the three datasets, the IDADP, the PlantVillage dataset, and the Tianjin Academy of Agricultural Sciences dataset, is 5336, 5963, and 7525, respectively. The five network models are tested on the three datasets. The experimental results are shown in Table 2.

Table 2 presents the experimental results of the average classification accuracy of the five network models on the three datasets. G-ResNet50 achieves higher classification results on the Tianjin Academy of Agricultural Sciences and the PlantVillage datasets, which is 2.3% higher than the IDADP. On the Tianjin Academy of Agricultural Sciences dataset, G-ResNet50 achieved the best classification rate higher than the other four network

models, the classification accuracy of G-ResNet50 on the Tianjin Academy of Agricultural Sciences dataset is only 1.2% higher than the PlantVillage dataset, the main reason is that the Tianjin Academy of Agricultural Sciences dataset has a diversified and large number of image enhancements, and the image background of the IDADP is too complicated.

4. Conclusions

Compared with the manual diagnosis of strawberry diseases, the computer vision method of identifying strawberry diseases has the advantages of low cost, high accuracy, and short time delay. Transfer learning and focal loss function are introduced in this network model, and the optimized G-ResNet50 is proposed based on deep residual network in this study.

The G-ResNet50 is trained on the strawberry disease dataset, and compared with the ResNet50, VGG16, InceptionV3, and MobileNetV2, the following conclusions are drawn:

Compared with the other 4 network models, the G-ResNet50 has faster convergence speed and higher generalization ability; the G-ResNet50 can perform deep training, which can alleviate the problem of performance degradation with the number of layers increases; the average recognition accuracy rate of the G-ResNet50 reached 98.67%, and the false recognition rate is lower than the other four networks. According to the evaluation index of the model, it can be found that the G-ResNet50 has better robustness and can be used for strawberry disease image recognition.

In the method of G-ResNet50, focus loss functions, transfer learning, and optimizing model parameters are introduced to improve the accuracy of model recognition. The method this study used still has some limitations; in the next research, the improvement of model [37, 38] performance by improving the relationship between the model network structure and optimization of model parameters will be considered, and new network models and optimization methods will be considered. So, in the future work, we will explore the new model network structure and optimization of model parameters for the identification of strawberry diseases.

Data Availability

All data included in this study are available upon request by contact with the corresponding author.

Consent

Not applicable.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

W.X. designed methodology; W.X. provided software; W.X. wrote the original manuscript and prepared the draft; W.X. wrote, reviewed, and edited the manuscript; Z.Y. supervised the study; and all authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded by the Tianjin Education Commission Scientific Research Project: Research on the Development of Collaborative Innovation under the Background of the Integration of Production and Education in Colleges (2018SK094).

References

- [1] H. F. Pardede, E. Suryawati, V. Zilvan, A. Ramdan, R. B. S. Kusumo, and A. Heryana, "Plant diseases detection with low resolution data using nested skip connections," *Journal of Big Data*, vol. 57, p. 7, 2020.
- [2] A. Jahanbakhshi, M. Momeny, M. Mahmoudi, and Y.-D. Zhang, "Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks," *Scientia Horticulturae*, vol. 263, Article ID 109133, 2020.
- [3] X. Xie, Y. Ma, B. Liu, J. He, S. Li, and H. Wang, "A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks," *Frontiers of Plant Science*, vol. 11, 2020.
- [4] G. L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, "Deep learning for plant identification using vein morphological patterns," *Computers and Electronics in Agriculture*, vol. 127, pp. 418–424, 2016.
- [5] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: classification and symptoms visualization," *Applied Artificial Intelligence*, vol. 31, 2017.
- [6] E. Fujita, Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, "Basic investigation on a robust and practical plant diagnostic system," in *Proceedings of the 2016 15th IEEE International Conference Mechanical Learning Application ICMLA*, vol. 2017, Anaheim, CA, USA, December 2016.
- [7] S. Ghosal, D. Blystone, A. K. Singh, and B. Ganapathysubramanian, "An explainable deep machine vision framework for plant stress phenotyping," *Proceedings of the National Academy of Sciences*, vol. 115, pp. 4613–4618, 2018.
- [8] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudiger, J. T. Heng, and U. R. Acharaya, "Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images," *Information Sciences*, vol. 441, pp. 41–49, 2018.
- [9] S. Yadav, N. Sengar, A. Singh, A. Singh, and M. K. Dutta, "Identification of disease using deep learning and evaluation of bacteriosis in peach leaf," *Ecological Informatics*, vol. 61, Article ID 101247, 2021.
- [10] A. K. Rangarajan, R. Purushothaman, and A. Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," *Procedia Computer Science*, vol. 133, pp. 1040–1047, 2018.
- [11] N. Zhu, X. Liu, Z. Liu et al., "Deep learning for smart agriculture: concepts, tools, applications, and opportunities," *International Journal of Agricultural and Biological Engineering*, vol. 11, no. 4, pp. 32–44, 2018.

- [12] V. Sravan, K. Swaraj, K. Meenakshi, and P. Kora, "A deep learning based crop disease classification using transfer learning," *Materials Today Proceedings*, vol. 6, 2021.
- [13] M. B. Tahir, M. A. Khan, K. Javed et al., "Recognition of apple leaf diseases using deep Learning and variances-controlled features reduction," *Microprocessors and Microsystems*, vol. 141, pp. 1–24, Article ID 104027, 2021.
- [14] J. Shin, Y. K. Chang, B. Heung, T. Nguyen-Quang, G. W. Price, and A. Al-Mallahi, "A deep learning approach for RGB image-based powdery mildew disease detection on strawberry leaves," *Computers and Electronics in Agriculture*, vol. 183, Article ID 106042, 2021.
- [15] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep neural networks with transfer learning in millet crop images," *Computers in Industry*, vol. 108, pp. 115–120, 2019.
- [16] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, and R. T. Hurst, "Convolutional neural networks for medical image analysis: full training or fine tuning?" *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.
- [17] A. Fotsokamgaguy, T. Akram, B. Laurent, S. R. Naqvi, M. MbomAlex, and N. Muhammad, "A deep heterogeneous feature fusion approach for automatic land-use classification," *Information Scientist*, vol. 467, no. 10, pp. 199–218, 2018.
- [18] Y. G. A. I. Shahin, K. M. Amin, and A. A. Sharawi, "White blood cells identification system based on convolutional deep neural learning networks," *Computer Methods and Programs in Biomedicine*, vol. 168, pp. 69–80, 2019.
- [19] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Frontiers of Plant Science*, vol. 8, pp. 1–7, 2017.
- [20] L. X. Boa Sorte, C. Toledo Ferraz, and F. Fambrini, "Coffee leaf disease recognition based on deep learning and texture attributes," *Procedia Computer Science*, vol. 159, pp. 135–144, 2019.
- [21] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm," *Information Processing in Agriculture*, vol. 7, pp. 249–260, 2020.
- [22] G. A. B. Jayme, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.
- [23] A. Z. da Costa, H. E. H. Figueroa, and J. A. Fracarolli, "Computer vision based detection of external defects on tomatoes using deep learning," *Biosystems Engineering*, vol. 190, pp. 131–144, 2020.
- [24] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.
- [25] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A Robust deep learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, p. 2022, 2017.
- [26] S. Pandiyan, M. Ashwin, R. Manikandan, K. M. Karthick Raghunath, and G. R. Anantha Raman, "Heterogeneous Internet of things organization predictive analysis platform for apple leaf diseases recognition," *Computer Communications*, vol. 154, pp. 99–110, 2020.
- [27] S. S. Chouhan, A. Kaul, U. P. Singh, and S. Jain, "Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: an automatic approach towards plant pathology," *IEEE Access*, vol. 6, pp. 8852–8863, 2018.
- [28] D. Oppenheim and G. Shani, "Potato disease classification using convolution neural networks," *Adv. Anim. Biosci.* vol. 8, p. 244, 2017.
- [29] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers of Plant Science*, vol. 7, pp. 1–10, 2016.
- [30] J. Boulent, S. Foucher, J. Théau, and P.-L. St-Charles, "Convolutional neural networks for the automatic identification of plant diseases," *Frontiers of Plant Science*, vol. 10, p. 941, 2019.
- [31] S. Qiao, Q. Wang, J. Zhang, and Z. Pei, "Detection and classification of early decay on blueberry based on improved deep residual 3D convolutional neural network in hyper-spectral images," *Scientific Programming*, vol. 2020, no. 4, 12 pages, 2020.
- [32] Z. Qingchen, B. Changchuan, L. C. Zhuo et al., "A GPU-based residual network for medical image classification in smart medicine - ScienceDirect," *Information Sciences*, vol. 536, pp. 91–100, 2020.
- [33] K. C. Lee, S. S. Roy, P. Samui, and V. Kumar, Eds., *Data Analytics in Biomedical Engineering and Healthcare*, Academic Press, Cambridge, MA, USA, 2020.
- [34] C. Hou, J. Zhuang, Yu Tang et al., "Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation," *Journal of Agriculture and Food Research*, vol. 5, Article ID 100154, 2021.
- [35] B. B. Traore, B. Kamsu-Foguem, and F. Tangara, "Deep convolution neural network for image recognition," *Ecological Informatics*, vol. 48, pp. 257–268, 2018.
- [36] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, p. 22, 2021.
- [37] K. Hu, Z. Zhang, and X. Niu, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, 2018.
- [38] Singkek Sci-Tech, "Agricultural Disease," 2018.