






# An apple leaf disease identification model for safeguarding apple food safety

Song LIU<sup>1</sup> , Haoran BAI<sup>1\*</sup> , Fengmei LI<sup>2\*</sup> , Dongwei WANG<sup>1</sup>, Yuhui ZHENG<sup>1</sup>,  
Qiupeng JIANG<sup>1</sup>, Fengbo SUN<sup>3</sup>

## Abstract

Apples are the fourth most produced fruit in the world, so it is important to safeguard them from disease damage. Although there are many deep learning-based plant disease detection models, the existing deep learning networks have complex structures and require large amounts of computational resources for support. Lightweight networks such as MobileNet and ShuffleNet designed for mobile devices could achieve better classification and faster recognition on public datasets, but their accuracy rates are not yet up to the requirements of practical applications. To address these problems, we constructed an improved apple leaf disease recognition algorithm based on MobileNetV2. We used feature reuse to improve the network structure, added a dense connection structure to the inverse residual module, and an ECA-Net attention module to increase the model's focus on diseased regions. We trained the improved model on the network on a dataset expanded by a generative adversarial network. The results showed that the improved model had a smaller number of model parameters, only 3.3 M, and a higher accuracy rate of 96.23% compared to Resnet50, ShuffleNet, and MobileNet models. The improved model had only 0.34 M addition in the number of parameters compared with MobileNet-V2, and had a 2.2% improvement in accuracy.

**Keywords:** apple leaf disease identification; MobileNetV2; data enhancement; attentional mechanisms; deep learning.

**Practical Application:** Identification of apple leaf diseases.

## 1 Introduction

Food safety is a problem that needs to be paid attention to. If people eat unsafe food, it will be a threat to their health. Excessive use of chemical fertilizers and pesticides is one of the most important factors that lead to food safety problems. So, if we can control plant diseases early, we can protect the food safety of agricultural products.

Because food safety is very worthy of attention, there have been many scholars in the world from a variety of aspects working to improve the research on food safety. Pires Martins et al. (2022) studied the vibriosis and its impact on microbiological food safety. Zhao & Talha (2022) studied the Evaluation of food safety problems based on the fuzzy comprehensive analysis method. The notion of a fuzzy expert system is explored in-depth, along with its rule-base and set membership functions.

Apples are the fourth largest fruit in the world in terms of total production and have a large share of the international market. The total world production of apples in 2014 was 84 million tons (Musacchi & Serra, 2018). The apple growing process could be threatened by diseases, and the appearance of apple diseases could affect the yield and quality of apples. Failure to spot the disease could lead to reduced yields or even extinction and the use of heavy pesticides. In traditional farming, apple diseases are usually monitored and identified manually. This method is not only labor-intensive but also overly dependent

on human experience. Therefore, we need a new technology for apple disease detection, that can be used to ensure the yield and quality safety of apples, which is important to ensure the food safety of apples.

With the continuous development of technology, the ability and intelligence of computers to process information are also improving, and neural network technology has also been greatly developed. Neural network technology has been used in different fields, such as food technology (Xu et al., 2022) and automatic control (Bai et al., 2022).

Krizhevsky et al. (2017) proposed AlexNet in 2012 and won the ImageNet classification competition. AlexNet demonstrated the promising potential of convolutional neural networks in the field of image processing and promoted the development of image recognition. After that, with the emergence of deep learning networks such as VGGNet (Simonyan & Zisserman, 2014), GoogleNet (Szegedy et al., 2015), and ResNet (He et al., 2016), and the improvement of computer hardware performance, the accuracy and application fields of image recognition had become more and more extensive. Nowadays, image recognition technology has been applied to various industries of social life, such as medical (Chang et al., 2019), transportation (Haghighat et al., 2020), and security (Tariq et al., 2020). Image

Received 15 Oct., 2022

Accepted 03 Dec., 2022

<sup>1</sup>College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao, China

<sup>2</sup>College of Food Science and Engineering, Qingdao Agricultural University, Qingdao, China

<sup>3</sup>China Zhongxin Construction Engineering Co. Ltd., Qingdao, China

\*Corresponding author: baihaoran111@126.com; lifm1114@126.com

recognition technology based on deep learning has been widely studied in the field of plant disease identification.

In 2019, Fang et al. (2019) proposed an apple leaf disease identification model based on the VGG16 model, which improved the accuracy of apple disease leaf classification. In 2021, Song et al. (2021) implemented an apple disease recognition model based on a small-scale dataset and achieved an accuracy of 98.5% on the dataset. In 2021, Chen & Yu (2022) conducted research on food safety sampling inspection system based on deep learning. According to their study results, deep learning outperforms other approaches. In 2022, Mahamudul Hashan et al. (2022) proposed a multilayer convolutional neural network MCNN to classify three apple leaf diseases, and the experimental results showed that the model achieved an accuracy of 98.4%. In 2022, Chu et al. (2022) proposed a Chenpi-year recognition method based on deep learning. They used data-enhanced dataset and an improved ResNet50 model to accurately identify the year of tangerine peels.

The existing problem is that with the emergence of deeper and more complex deep learning networks, it leads to an increasing number of parameters and computation of deep models (Lin et al., 2017), which leads to difficulties in model training, long prediction times, and high hardware performance requirements (Russakovsky et al., 2015). In recent years, the research of experts and scholars has turned to how to improve model efficiency, and many lightweight neural networks have started to appear, such as SqueezeNet (Iandola et al., 2016), Xception (Chollet, 2017), MobileNet (Howard et al., 2017), ShuffleNet (Zhang et al., 2018), etc. The structure of such lightweight deep network models was simpler and requires fewer computational resources than traditional deep networks, which was more beneficial for practical production applications.

In this paper, we proposed an improved network ECA-DCMobileNet based on Mobilenet-V2 (Sandler et al., 2018), introduced the dense connection structure of DenseNet (Huang et al., 2017) and incorporated the ECA (Wang et al., 2020) channel attention module. ECA-DCMobileNet performed better than traditional deep networks with fewer parameters and optimal accuracy on a self-built apple leaf disease dataset augmented with GAN networks.

## 2 Method

The ECA-DCMobileNet proposed in this paper is based on the MobileNet-V2 (Sandler et al., 2018) network structure, and the densely connected structure of the ECA (Wang et al., 2020) channel attention module and DenseNet (Huang et al., 2017) was added to the step-1 antiresidual module to form a new densely connected antiresidual module. We replaced all the inverse residual modules of step size 1 in MobileNet-V2. The improved network structure is shown in Figure 1.

### 2.1 MobileNet-V2

The Google team proposed MobileNet (Howard et al., 2017) in 2017, which was a lightweight convolutional neural network specifically designed for mobile and embedded devices. The team then proposed the MobileNet-V2 (Sandler et al., 2018) version in 2018. The MobileNet family of networks could effectively reduce the number of parameters and computation of the model with only a small loss of accuracy.

MobileNet splits the standard convolution into depth-wise convolution and point-wise convolution, where depth-wise convolution performs the convolution operation on a single channel of the input image. Point-wise convolution performs

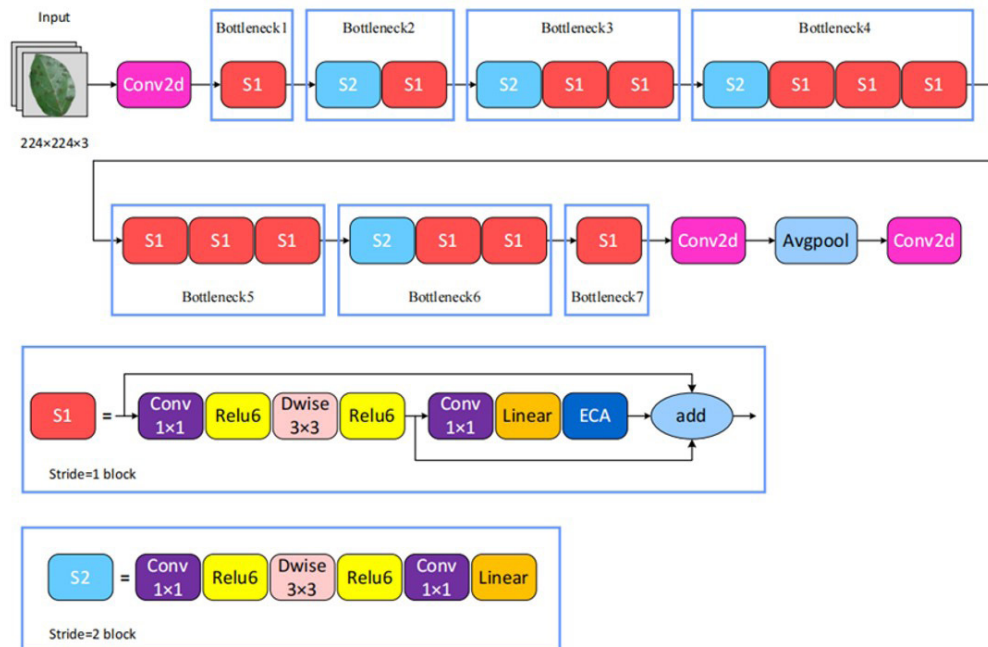


Figure 1. Improved network structure.

the calculation of a linear combination of the input image to generate a new feature map, as shown in Figure 2.

MobileNet-V2 introduces an inverse residual network based on MobileNet’s deep-separated convolution and uses two  $1 \times 1$  point-by-point convolutions in each block to up-dimension and down-dimension the data. It uses deep convolution to process the image data with 6-fold up-dimensioning and uses linear activation instead of the original Relu6 activation function after the down-dimensional convolution in each block to mitigate the problem of information loss. This is shown in Figure 3. This structure is the opposite of the residual network structure of ResNet (He et al., 2016) where dimensionality is reduced before

dimensionality is increased, and is therefore called the inverse residual structure. The overall structure of MobileNet-V2 is shown in Figure 4 and Table 1.

### 2.2 Attention module ECA-Net

ECA-Net (Wang et al., 2020) is a novel channel attention convolutional neural network with an improved channel attention module based on SENet (Wei et al., 2017), which proposes a lightweight module ECA module. It proposes a dimensionless local cross-channel interaction strategy and a channel dimension function that adaptively determines the size of the one-dimensional convolutional kernel, which both enable this network attention

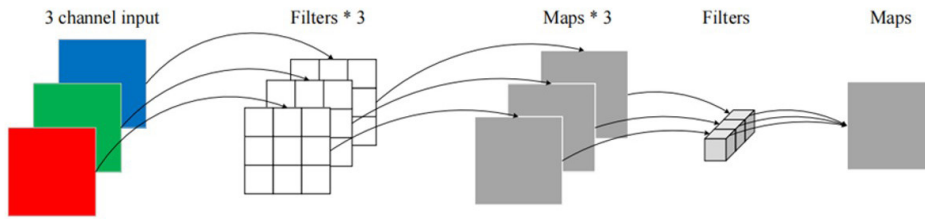


Figure 2. Deep Separation Convolution.

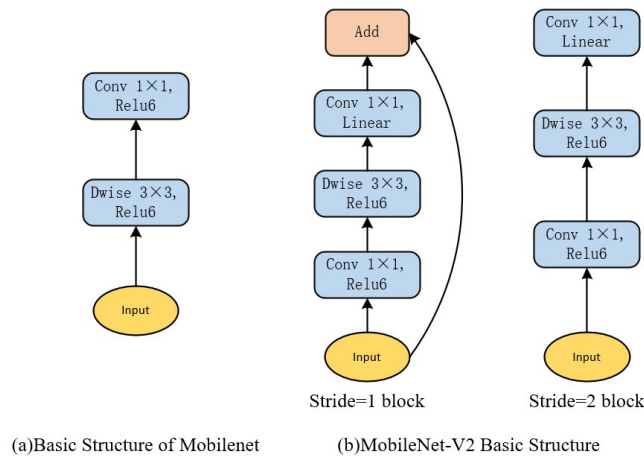


Figure 3. Basic structure of MobileNet and Mobilenet-V2. (a) Basic structure of MobileNet. (b) MobileNet-V2 basic structure.

Table 1. MobileNet-V2 overall structure.

Input	Operator	t	c	n	s
$224^2 \times 3$	Conv2d	-	32	1	2
$112^2 \times 32$	Bottleneck	1	16	1	1
$112^2 \times 16$	Bottleneck	6	24	2	2
$56^2 \times 24$	Bottleneck	6	32	3	2
$28^2 \times 32$	Bottleneck	6	64	4	2
$28^2 \times 64$	Bottleneck	6	96	3	1
$14^2 \times 96$	Bottleneck	6	160	3	2
$7^2 \times 160$	Bottleneck	6	320	1	1
$7^2 \times 320$	Conv2d $1 \times 1$	-	1280	1	1
$7^2 \times 1280$	Avgpool $7 \times 7$	-	-	1	-
$1 \times 1 \times k$	Conv2d $1 \times 1$	-	k	-	-

Note: t denotes the expansion multiple, c denotes the number of output channels, n denotes the number of repetitions, and s denotes the step size. In Bottleneck with step size 2, only the first repetition step size is 2, and the subsequent repetition steps are all 1.

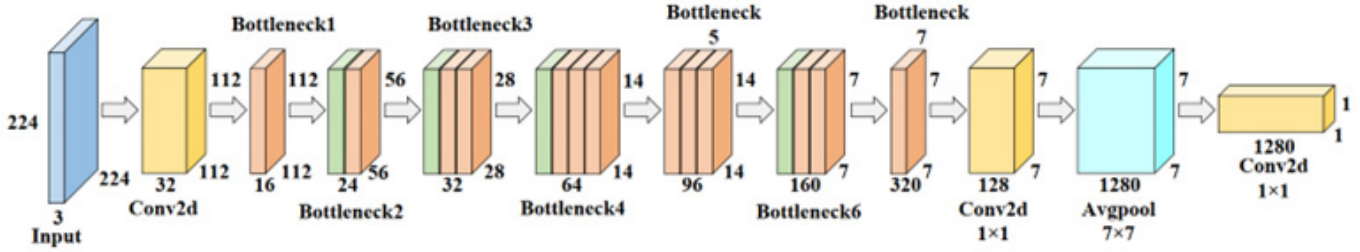


Figure 4. Schematic diagram of Mobilenet-V2 network structure.

mechanism to gain performance improvement and reduce the complexity of the model at the same time. The structure diagram of the ECA module is shown in Figure 5.

Han et al. (2021) proposed a method for insulator detection and damage identification in 2021, which embedded the ECA-Net attention module in Tiny-YOLOv4. The results showed that the inclusion of the ECA-Net module reduced the model complexity, ensured the detection accuracy of the model, and guaranteed the effectiveness of the model for multi-class and multi-scale objects.

In MobileNet-V2, the base module is the inverse residual module, which is similar to the residual module. We add the ECA channel attention module after the  $1 \times 1$  reduced convolution of the inverse residual module with step size equal to 1, and improve the inverse residual module by taking advantage of the channel attention mechanism of the ECA module. The 1D convolution is adjusted in ECA by adaptive convolution kernel size in the channel attention module instead of the fully connected layer (Wang et al., 2020), which enables the anti-disparity network to focus the recognition region on the difference of image, thus improving the recognition performance of the network. A schematic diagram of the location of the ECA module in the improved network is shown in Figure 6.

2.3 Dense connection structure

The dense connection structure is used in DenseNet (Huang et al., 2017), i.e., in a dense block, which connects each layer to every other layer in a feed-forward fashion.

Dense connectivity helps to alleviate the gradient disappearance problem. The gradient disappearance problem is more likely to occur in deeper networks because it is caused by multiple passes of input and gradient information between different layers. In a densely connected structure, the input of each layer is directly connected to the previous output, thus mitigating the gradient vanishing phenomenon. In addition, the dense connection has the effect of regularization, which is useful for overfitting (Huang et al., 2017). The structure of the dense block is shown in Figure 7.

In 2020, Cui et al. (2020) used a dense connectivity structure in the decoder subnet to fuse multiscale information from each layer to enhance the representation of features. This improvement was used to enhance the performance of the semantic segmentation model. After experiments on remote sensing images datasets, it is shown that the dense connection

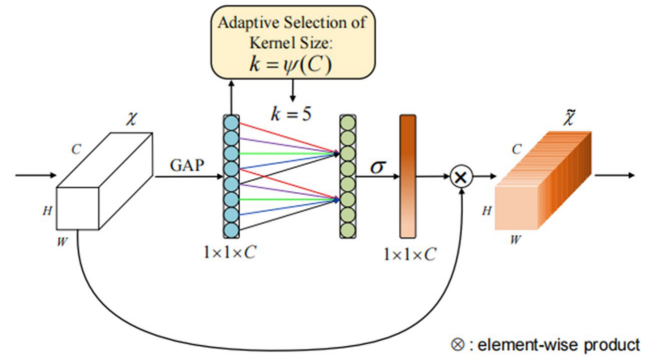


Figure 5. ECA module structure diagram

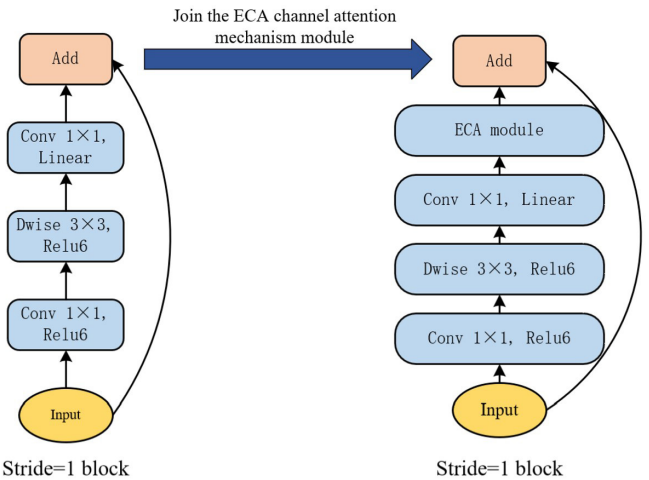


Figure 6. Add the ECA channel attention module.

structure could effectively solve the problem of insufficient data annotation. The improved model had better performance and more accurate segmentation results.

We used the dense connection structure of dense block in the ECA-DCMobileNet-V2 module with step size 1. In the original inverse residual module shown in Figure 1(a), the output  $XI\_5$  is the sum of the original input  $XI\_0$  and  $XI\_0$  of this inverse residual module after the nonlinear operation of this module and the two parts of the output  $XI\_4$ , expressed as Equation 1:

$$XI\_5 = XI\_4 + XI\_0 = NI\_1234(XI\_0) + XI\_0 \tag{1}$$

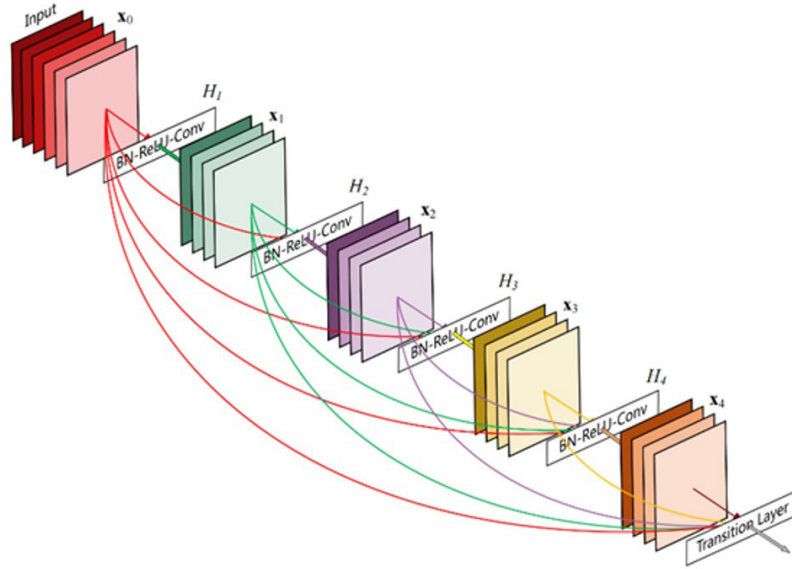


Figure 7. Dense block structure diagram.

NI\_1234(\*) represents all operations performed on the input inverse residual module data. When performing feature map summation, it is required that the feature maps have the same size, so XI\_4 should have the same number of rows, columns, and layers as the input image XI\_0.

We added a branch connecting XI\_2 to XI\_4 on top of the inverse residual module with the ECA module, as shown in Figure 8(b). We obtained the base module of ECA-DCMobileNet-V2. Theoretically, there was no limit to the number of branches that could be added, but to improve feature utilization and limit the amount of computation, we only used the output at XI\_2. Because the output at XI\_2 was located on a higher dimension and could yield more information.

As can be seen in Figure 8, in the densely connected inverse residual module, the input XI\_0 and the feature map of the output XI\_2 after deep convolution were added to the final feature map processed by the ECA module to obtain the final output. It should be noted that the feature maps should be of the same size when added, and we used a 1 × 1 convolution with a step size of 2 for the transformation.

The input relations for the densely connected inverse residual module in Figure 1(b) are as follows, expressed as Equations 2-6:

$$XI_1 = NI_1(XI_0) \tag{2}$$

$$XI_2 = NI_2(XI_1) \tag{3}$$

$$XI_3 = LI_3(XI_2) \tag{4}$$

$$XI_4 = NI_4(XI_3) \tag{5}$$

$$XI_5 = XI_4 + XI_2 + XI_0 = NI_4(LI_3(NI_2(NI_1(XI_0)))) + NI_2(NI_1(XI_0)) + XI_0 \tag{6}$$

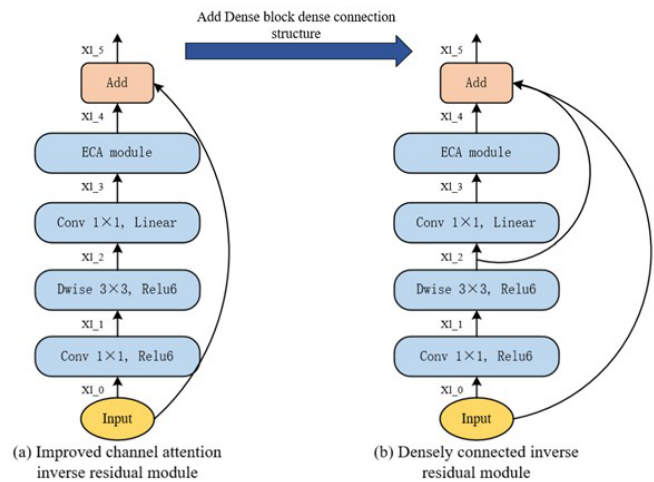


Figure 8. Add dense connection structure to the original channel attention improved inverse residual module. (a) Improved channel attention inverse residual module. (b) Densely connected inverse residual module.

XI\_\* denotes the feature map corresponding to each step, NI\_\*(-) denotes the corresponding nonlinear operation, and LI\_\* denotes the corresponding linear operation. XI\_5 denotes the output of the entire densely connected inverse residual module, which is obtained by summing the module input, the output of the depth convolution part, and the output of the ECA channel attention module.

### 3 Results and analysis

The hardware and software environments used in this experiment are shown in Table 2.

### 3.1 Dataset preparation

#### Image dataset

The object of this thesis was the images of apple leaf diseases. We collected images of apple leaves with eight different symptoms. The dataset included spotted leaf drop, powdery mildew, brown spot, gray spot, rust, blight rot, anthracnose blight, and healthy leaves. 700 images of leaf diseases were collected, ranging from 60 to 140 images for each symptom. The images were taken under natural light conditions, and the apple variety was Red Fuji. The images were collected in Qixia, Shandong Province, China (120°64'499" E, 37°31'816" N). Examples of apple leaf symptoms are shown in Figure 9.

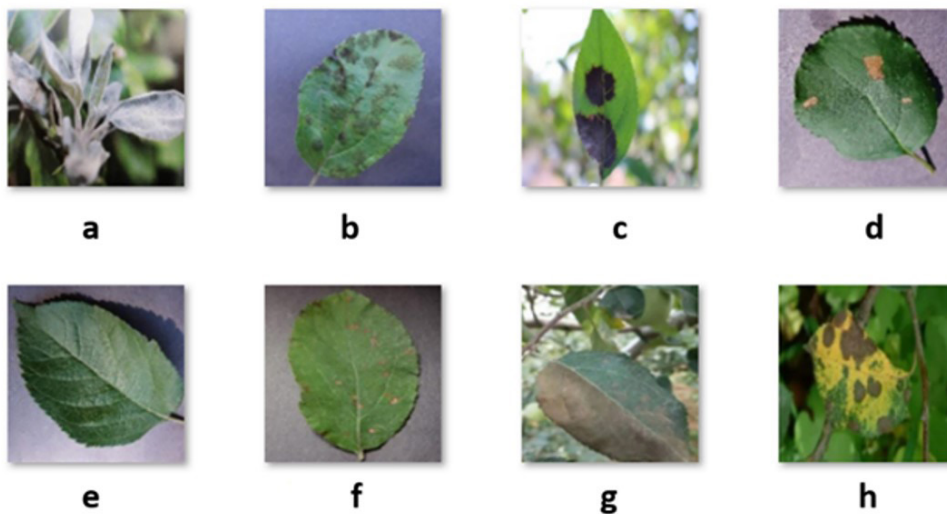
#### Data enhancement

When training for deep learning, the more images, the better the training effect would be and the stronger the generalization ability of the obtained model would be. Therefore, enough images were required for training. The 700 images we prepared were not enough to meet the training requirements, so we used data enhancement to expand the dataset and raise the quality of the data to improve the generalization ability and robustness of the model.

Geometric transformation-based data enhancement methods (traditional data enhancement) mostly used twisting, stretching,

**Table 2.** Operating environment-related parameters.

Name	Versions
CPU	Intel(R) core(TM) i7-10750H CPU @2.60GHZ
GPU	NVIDA GEFORCE GTX 1650 4G
Memory	16GB
Operating System	Microsoft Windows 11 Professional (64bit)
Development framework	Tensorflow \ Keras
Programming language	Pyhton



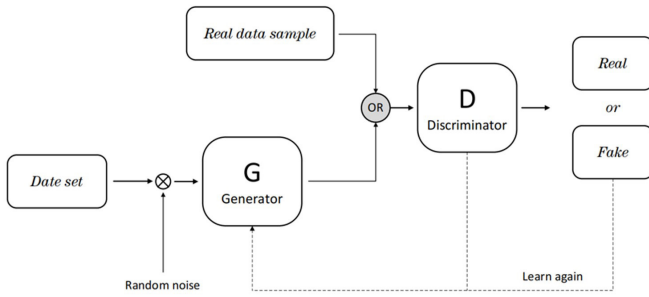
**Figure 9.** Example of 8 kinds of pictures in self-built apple disease dataset. a: powdery mildew. b: spotted leaf drops. c: brown spot. d: gray spot. e: healthy leaves. f: rust g: blight rot. h: anthracnose blight.

erasing, noise, etc. to expand the data. However, geometric transformation methods could change and remove some of the original features of the image, which had an impact on the accuracy of the training results. Therefore, we used generative adversarial networks (Goodfellow et al., 2020) to expand the dataset.

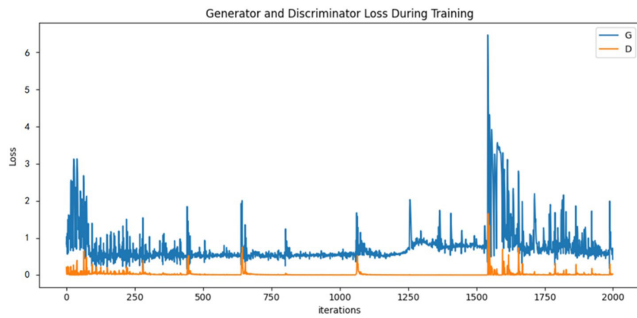
Generative adversarial network was a learning method that used the idea of “adversarial” to generate models by Goodfellow et al. It was essentially a deep learning model and was one of the most used methods in unsupervised data enhancement. The two modules of the generative adversarial network, generative and discriminative, confront and learn from each other as shown in Figure 10. Finally, the generative adversarial network could generate images that could be faked to achieve a great expansion of the dataset with a small loss of image features.

For training with GAN, we used the Adam optimization algorithm of the stochastic gradient descent method, which could use “inertia” to get the model out of the local optimal state, so that the model could be close to the global optimal. We set the Adam optimizer parameter to 0.5 and the learning rate of the optimizer to 0.0003. We set the noise dimension to 100, trained the discriminator once every 1 batch, and generated the model once every 5 batches. Taking apple rust as an example, the loss curve is shown in Figure 11, and the generated images in each stage are shown in Figure 12.

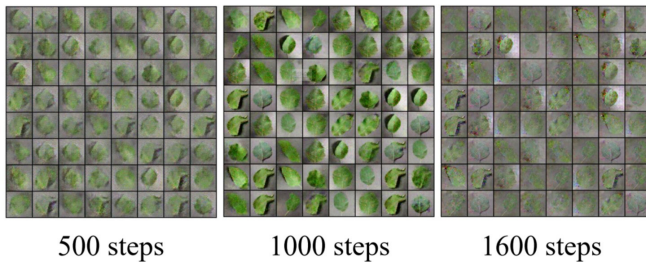
According to the loss curve shown in Figure 11 and the generated images at each stage shown in Figure 12. The GAN generated the network loss function and the discriminative network loss function no longer fluctuated drastically at 500 times, and the quality of the generated images gradually rose at this time. When the number of training times reached about 1250, the loss value started to rise, and the quality of the generated images started to decline. The generated images were the clearest at 1250 training times, and the blurred and broken images already appear when the number of training times reaches 1600. The example of the generated images is shown in Figure 12.



**Figure 10.** Generating adversarial network process.



**Figure 11.** Loss function curve of GAN.



**Figure 12.** Use the images generated by the generative adversarial network at each stage.

After generating an adversarial network to expand the dataset, we obtained a total of 11,100 images, as shown in Table 3. Among them, the number of healthy leaves and spots of deciduous disease after expansion larger than the number of other species. In order to avoid a skewed distribution in the model training due to the difference in the number of samples, we reduced samples to 1500 to make the number of samples in each category approximately the same. The size of all images was uniformly set to 224\*224, and the data set was divided into training and validation sets in the ratio of 8:2.

### 3.2 Analysis of experimental results

On the self-built apple leaf disease dataset. Table 4 shows the accuracy, recall, and number of model parameters of five deep networks, including improving network. As can be seen from the table, the improved model had the highest accuracy and recall, reaching 96.2% and 99.5%, respectively. In terms of parameters, the number of parameters of the improved model was

**Table 3.** Distribution of data sets.

Type	Number of original images	Number of extended images
Powdery mildew	60	1200
Spot leaf	100	1500
Cercospora leaf spot	90	1500
Grey blight	70	1300
Healthy	140	1500
Rust	80	1400
Epidemic	70	1300
Colletotrichum blight	90	1400
Aggregate	700	11100

**Table 4.** Performance comparison of different networks.

Name	Accuracy	Recall	Parameters /M
AlexNet	89.2%	92.6%	105
VGG16	92.8%	96.8%	138
ResNet50	95.6%	97.9%	25.6
ShuffleNet	94.5%	98.6%	-
MobileNet-V2	94.7%	99.4%	2.96
New model	96.2%	99.5%	3.5

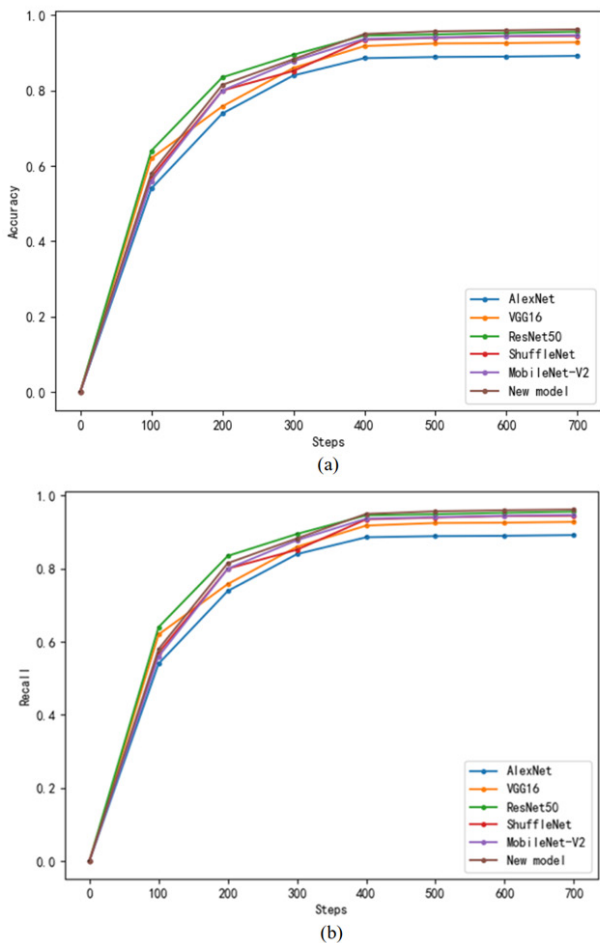
only 0.54 M higher than that of MobileNet-V2. All these metrics illustrated the superiority of ECA-DCMobileNet and proved that it had better results on our self-built apple leaf disease dataset. The variation curves of accuracy and recall during training of the six models are shown in Figure 13.

As shown in Table 5, our model performed well in terms of classification accuracy for the eight kinds of images in the self-built dataset. We achieved 95.6% and 96.1% accuracy for gray spot and rust diseases. In terms of accuracy, gray spot disease had the lowest of 94.5%. In terms of recall, rust disease had the lowest of 92.1%. In terms of F1-score, gray spot and blight rot were 93.9% and 92.6%, lower than others, respectively. The F1-score of other categories were higher than 95%. The experimental results shown that our improved model performs well in apple leaf disease identification.

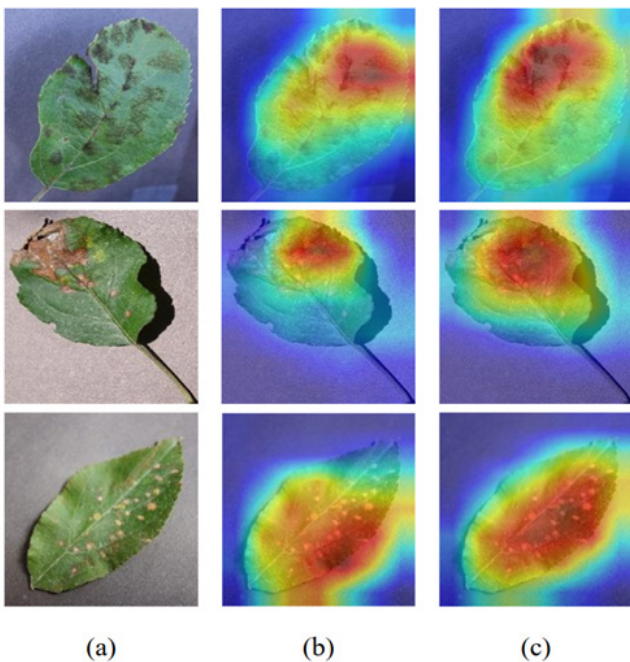
### 3.3 Visualization results analysis

To further verify the effectiveness of the improved network for apple leaf disease identification, we used Class Activation Mapping (CAM) for image visualization and obtained the heat map of MobileNet-V2 with the improved network ECA-DCMobileNet, as shown in Figure 14.

The original image of apple leaf disease is shown in Figure 14(a), the MobileNet-V2 visualized heat map in Figure 14(b), and the ECA-DCMobileNet visualized heat map in Figure 14(c). Through the comparison between Figure 14(b) and Figure 14(c), the improved model has significantly increased the attention to the disease feature concentration area on the leaves. It also pays less attention to other unrelated areas. The results demonstrated that the improved model had enhanced directionality in judging disease features and improved the recognition of apple leaf diseases.



**Figure 13.** Comparison of recall and accuracy of six models. (a) The accuracy of six models. (b) The recall of six models.



**Figure 14.** Class activation visualization heat map. (a) Original image. (b) The heat map of MobileNet-V2. (c) The heat map of ECA-DCMobileNet.

**Table 5.** Detailed classification results of various types of diseases.

Name	Accuracy	Precision	Recall	F1-score
Powdery mildew	98.2%	100.0%	100.0%	99.5%
Spot leaf	97.9%	97.1%	97.5%	99.1%
Cercospora leaf spot	97.5%	97.3%	98.1%	99.3%
Grey blight	95.6%	94.5%	92.9%	93.9%
Healthy	98.8%	100.0%	99.5%	99.8%
Rust	96.1%	94.8%	92.1%	92.6%
Epidemic	96.8%	98.6%	98.6%	98.5%
Colletotrichum blight	97.7%	98.3%	97.8%	98.2%

### 4 Conclusion

In this paper, we proposed an improved deep network architecture based on MobileNet-V2. We introduced the idea of dense connection to improve the feature reuse ability of the network. The ECA attention module was added to increase the attention of the network to the feature regions. We used GAN for data enhancement and built our own dataset of apple leaf diseases containing eight categories. Finally, we used the ECA-DCMobileNet to conduct experiments on the self-built dataset. We conducted comparative experiments on model performance:

- (1) We used our model to compare performance with AlexNet, VGG16, ResNet50, ShuffleNet, and MobileNet-V2. The experiments shown that the accuracy and recall rate of the new model were the best, and its parameter quantity was also the lowest.
- (2) We used the new model to test eight kinds of images in the self-built data set. The classification accuracy of the new model for eight kinds of images was greater than 95%. The model had good generalization ability.
- (3) We used Class Activation Mapping (CAM) for image visualization and compared the heat maps of MobileNet-V2 and the improved network ECA-DCMobileNet. The new model paid more attention to the disease area on leaves, which proved that the new model had better feature extraction ability. It could provide a reference for the classification of apple leaf disease images.

The experimental results shown that the improved network achieved further improvement in classification accuracy while maintaining the number of parameters. Our model improved training and recognition efficiency. In the subsequent research, we will further improve the model from the practical application, reduce the network complexity, and apply it to target recognition.

The focus of our future work is as follow:

- (1) In order to enhance the detection ability of the model, so that it can identify apple leaves in a variety of environments, we need to collect more data set images.



- (2) The feature extraction network of the model can also be improved, so that the model of apple leaf disease feature extraction capability will be improved.
- (3) An apple leaf disease identification system needs to be built. We want to deploy our model and service website on the server so that people can login to the website to use the model to identify apple disease leaf images.

## Acknowledgements

This work was supported by Guangdong Province Key Field Research and Development Plan (Project no.2018B020241003-09), Shandong Provincial Natural Science Foundation (Project no. ZR2021ME018), Shandong Province Agricultural Machinery Equipment Research and Development Innovation Plan (Project no. 2018YZ002), National Peanut Industry Technology System (CARS-13-Mechanized Sowing and Field Management Positions) (Project no. S202010435013), Project of China Construction Center Construction Engineering Co., LTD (Project no. ZX&AZ02202200016001). Qingdao Agricultural University ideological and political education program (Project no. S1814).

## References

- Bai, H., Chu, Z., Wang, D., Bao, Y., Qin, L., Zheng, Y., & Li, F. (2022). Predictive control of microwave hot-air coupled drying model based on GWO-BP neural network. *Drying Technology*. In press. <http://dx.doi.org/10.1080/07373937.2022.2124262>.
- Chang, W. J., Chen, L. B., Hsu, C. H., Lin, C. P., & Yang, T. C. (2019). A deep learning-based intelligent medicine recognition system for chronic patients. *IEEE Access: Practical Innovations, Open Solutions*, 7, 44441-44458. <http://dx.doi.org/10.1109/ACCESS.2019.2908843>.
- Chen, T. C., & Yu, S. Y. (2022). Research on food safety sampling inspection system based on deep learning. *Food Science and Technology (Campinas)*, 42, e29121. <http://dx.doi.org/10.1590/fst.29121>.
- Chollet, F. (2017). Xception: deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1251-1258). New York: IEEE. <http://dx.doi.org/10.1109/CVPR.2017.195>.
- Chu, Z., Li, F., Wang, D., Xu, S., Gao, C., & Bai, H. (2022). Research on identification method of tangerine peel year based on deep learning. *Food Science and Technology (Campinas)*, 42, e64722. <http://dx.doi.org/10.1590/fst.64722>.
- Cui, B., Chen, X., & Lu, Y. (2020). Semantic segmentation of remote sensing images using transfer learning and deep convolutional neural network with dense connection. *IEEE Access: Practical Innovations, Open Solutions*, 8, 116744-116755. <http://dx.doi.org/10.1109/ACCESS.2020.3003914>.
- Fang, T., Chen, P., Zhang, J., & Wang, B. (2019, August). Identification of apple leaf diseases based on convolutional neural network. In *International Conference on Intelligent Computing* (pp. 553-564). Cham: Springer. [http://dx.doi.org/10.1007/978-3-030-26763-6\\_53](http://dx.doi.org/10.1007/978-3-030-26763-6_53).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. <http://dx.doi.org/10.1145/3422622>.
- Haghighat, A. K., Ravichandra-Mouli, V., Chakraborty, P., Esfandiari, Y., Arabi, S., & Sharma, A. (2020). Applications of deep learning in intelligent transportation systems. *Journal of Big Data Analytics in Transportation*, 2(2), 115-145. <http://dx.doi.org/10.1007/s42421-020-00020-1>.
- Han, G., He, M., Zhao, F., Xu, Z., Zhang, M., & Qin, L. (2021). Insulator detection and damage identification based on improved lightweight YOLOv4 network. *Energy Reports*, 7, 187-197. <http://dx.doi.org/10.1016/j.egy.2021.10.039>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770-778). Las Vegas: IEEE. <https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.90>.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv*. In press. <https://doi.org/10.48550/arXiv.1704.04861>.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700-4708). Las Vegas: IEEE. <https://doi.org/10.48550/arXiv.1608.06993>.
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. *arXiv*. In press. <https://doi.org/10.48550/arXiv.1602.07360>.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90. <http://dx.doi.org/10.1145/3065386>.
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988). Venice: IEEE. <https://doi.org/10.48550/arXiv.1708.02002>.
- Mahamudul Hashan, A., Md Rakib Ul Islam, R., & Avinash, K. (2022). Apple leaf disease classification using image dataset: a multilayer convolutional neural network approach. *Informatics and Automation*, 21(4), 710-728. <http://dx.doi.org/10.15622/ia.21.4.3>.
- Musacchi, S., & Serra, S. (2018). Apple fruit quality: overview on pre-harvest factors. *Scientia Horticulturae*, 234, 409-430. <http://dx.doi.org/10.1016/j.scienta.2017.12.057>.
- Pires Martins, V. G., Santos Nascimento, J., Silva Martins, F. M., & Ceotto Vigoder, H. (2022). Vibriosis and its impact on microbiological food safety. *Food Science and Technology (Campinas)*, 42, e65321. <http://dx.doi.org/10.1590/fst.65321>.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211-252. <http://dx.doi.org/10.1007/s11263-015-0816-y>.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520). Salt Lake City: IEEE. <https://doi.org/10.48550/arXiv.1801.04381>.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv*. In press. <https://doi.org/10.48550/arXiv.1409.1556>.
- Song, C., Wang, D., Bai, H., & Sun, W. (2021). Apple disease recognition based on small-scale data sets. *Applied Engineering in Agriculture*, 37(3), 481-490. <http://dx.doi.org/10.13031/aea.14187>.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Banhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer*

- Vision and Pattern Recognition* (pp. 1-9). Palo Alto: IEEE. <https://doi.ieeecomputersociety.org/10.1109/CVPR.2015.7298594>.
- Tariq, M. I., Memon, N. A., Ahmed, S., Tayyaba, S., Mushtaq, M. T., Mian, N. A., Imran, M., & Ashraf, M. W. (2020). A review of deep learning security and privacy defensive techniques. *Mobile Information Systems*, 2020, 1-18. <http://dx.doi.org/10.1155/2020/6535834>.
- Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). ECA-Net: efficient channel attention for deep convolutional neural networks. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11531-11539). Seattle: IEEE. <http://dx.doi.org/10.1109/CVPR42600.2020.01155>.
- Wei, X. S., Luo, J. H., Wu, J., & Zhou, Z. H. (2017). Selective convolutional descriptor aggregation for fine-grained image retrieval. *IEEE Transactions on Image Processing*, 26(6), 2868-2881. <http://dx.doi.org/10.1109/TIP.2017.2688133>. PMID:28368819.
- Xu, X., Ren, S., Wang, D., Ma, J., Yan, X., Guo, Y., Liu, X., & Pan, Y. (2022). Optimization of extraction of defatted walnut powder by ultrasonic assisted and artificial neural network. *Food Science and Technology (Campinas)*, 42, e53320. <http://dx.doi.org/10.1590/fst.53320>.
- Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: an extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6848-6856). Salt Lake City: IEEE. <http://dx.doi.org/10.1109/CVPR.2018.00716>.
- Zhao, Y., & Talha, M. (2022). Evaluation of food safety problems based on the fuzzy comprehensive analysis method. *Food Science and Technology (Campinas)*, 42, e47321. <http://dx.doi.org/10.1590/fst.47321>.