Intelligent Diagnosis and Prediction of Plant Diseases Based on Efficientnet_B3a

Jian Liu [1, *, #], Ruodi Huang [2, #], Xinru Ma [3, #]

[1] School of Control and Computer Engineering, North China Electric Power University, Baoding, China, 071000

[2] School of Internet, Anhui University, Hefei, China, 230039

[3] School of Information Engineering, Tianjin University of Commerce, Tianjin, China, 300133

* Corresponding Author Email: yc_2002_151@163.com

[#]These authors contributed equally.

Abstract. Crop diseases pose a significant threat to agricultural production, affecting plant growth, reducing yields, and ultimately hampering the agricultural economy. Therefore, accurate detection of plant diseases is paramount. Recent advancements in deep learning have demonstrated superior performance compared to traditional recognition methods in the realm of intelligent diagnosis. Motivated by this, the present study employs a multi-network model based on transfer learning to identify 39 distinct types of plant diseases. We innovatively select the Efficientnet_b3a network model for plant leaf disease recognition. The Efficientnet_b3a is used for the first time in plant detection. This model boasts the advantages of high accuracy and low network complexity. To further validate its performance, we utilize three established network models, namely Resnet50, Google Net, and Inception_V3, as benchmarks for comparison. To validate the effectiveness of the Efficientnet_b3a model, this paper employs the publicly available Plant Village dataset to conduct training, validation, and testing on four different network models for comparative analysis. The experimental results demonstrate that the Efficientnet_b3a model achieves a remarkable recognition accuracy of 99.13% in plant disease identification, outperforming traditional network models. Notably, it manages to reduce the complexity of the network and improve the efficiency. At the same time, it is used in the rapid identification of some complex scenes. Compared with the classical complex network, the method verified in this paper has advantages in the rapid diagnosis of plant diseases. Furthermore, the model exhibits excellent network complexity, making it suitable for plant leaf disease recognition in practical agricultural production. This provides an efficient and accurate solution for disease detection in agricultural production, thus exhibiting broad application prospects.

Keywords: Plant Disease Detection; Efficientnet_b3a; Transfer Learning.

1. Introduction

In the agricultural field, early detection of plant diseases is crucial for safeguarding crop yields and quality. Traditional detection methods often rely on manual observation, leading to extremely low efficiency and the possibility of missed detections[1]. At the same time, one of the challenges in plant disease detection lies in the diversity of disease types and the variability of disease severity. Deep learning-based image processing and object detection models, as advanced techniques, can efficiently and accurately detect multiple targets, thus becoming a potential method for plant disease detection. These methods can provide farmers with timely disease warnings, enabling them to take rapid and targeted prevention and control measures, thereby improving crop yields and quality. In image processing systems, digital signals exhibit rich two-dimensional characteristics and utilize multi-level signal processing techniques[2]. Digital image processing (DIP) refers to the process of converting images into digital form through a series of operations, to obtain better visualization effects or extract key features from the images[3].

In recent years, deep learning (DL) has developed rapidly, primarily used for modeling the complex relationships between data and features extracted from inputs. Because of this, deep learning now plays an extremely important role in various industries [4]. Compared with traditional techniques such as edge detection and filtering, deep learning technologies like Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) not only demonstrate excellent performance in handling large-scale and diverse data, achieving optimal results but also exhibit remarkable performance in fields such as image classification and object detection, driving the development of computer vision and pattern recognition [5].

Especially after the introduction of the Convolutional Neural Network (CNN) model, the application of DL in the field of image processing has been rapidly promoted [6]. Ma employed CNN in hyperspectral image detection research and applied it to the detection of multi-focus images [7]. Prakhar Bansal proposed a collection of deep learning models, including DenseNet121, EfficientNetB7, and Efficient Net Noisy Student. A detailed analysis of the performance of this model ensemble revealed an excellent accuracy of 96.25%, demonstrating the significant effect of model integration in improving performance [8]. Bincy Chellapandi utilized eight CNN models for plant disease detection research, including VGG16, VGG19, Resnet50, MobileNetV2, Inception ResnetV2, and DenseNet. This comprehensive research method helped me gain a deeper understanding of the advantages and limitations of various models in processing plant disease images [9]. Akshai KP adopted a comprehensive approach to plant disease recognition, combining CNN models with pre-trained models such as DenseNet, VGG, and Resnet. The experimental results achieved an accuracy of up to 98.27%. Additionally, the study found that DenseNet is capable of handling various complex features and maintaining high accuracy even on small-sample datasets [10]. These studies provide strong support for the application of deep learning in plant disease detection and serve as a reference for future research and applications.

Based on a large number of previous studies, we select the simple and efficient neural network model of Efficientnet_b3a. After modifying its network structure, the Plant Village dataset is used for training, validation, and testing. The model is also used to predict random disease problems. To compare the experimental results, this paper selects common classical network models such as Resnet50, Google Net, and Inception_V3 for comparative analysis. The research results show that the adopted network model has better accuracy and lower network complexity compared with the classic networks. This method ultimately achieves rapid detection of plant diseases and has a positive impact on the protection of crops.

2. Methodology

In this study, we employed the novel and efficient Efficientnet_b3a model to conduct intelligent diagnosis and prediction of crop diseases using the Plant Village dataset. To highlight the advantages of this network model, we also selected three classic network models, namely Resnet50, Google Net, and Inception_V3, for training, validation, and testing under the same dataset and classification conditions. As shown in Figure 1, this is the basic technical framework of our study. Firstly, we conducted certain data enhancement processing on the public dataset and trained it under different network models to obtain the optimal network model weights. Then, we used these weights to test and predict different plant diseases. The output dataset classification targets included 39 categories of 14 different species. This method can accurately extract respective features such as shape, color, and texture from different types of leaves, thereby providing more accurate results. Before training, we preprocessed the training and validation data, adjusting the image size to 224×224 (for Inception_V3 model training, the size was adjusted to 299×299) according to the input and output requirements of the network model. After certain processing, the data was converted into the Tensor format and standardized for image processing by PyTorch.

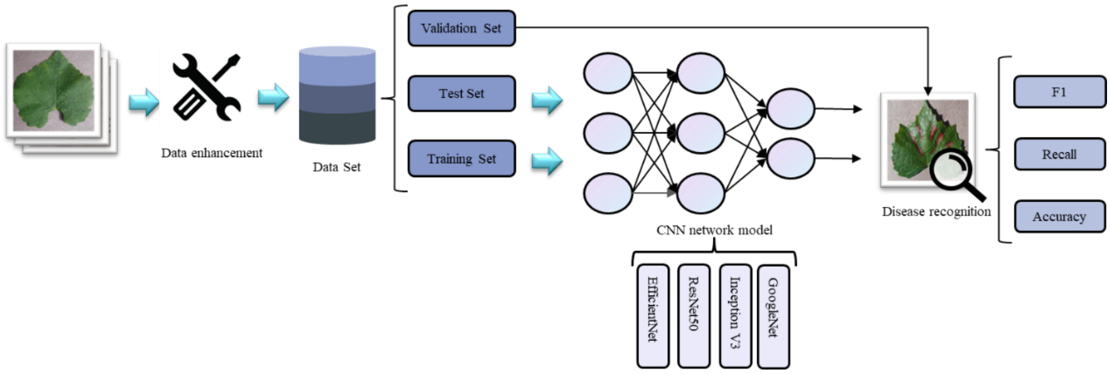

Figure 1: Flowchart

2.1. Dataset

The dataset for this study was obtained from Plant Village, which comprises 39 plant lesion category categories and 55,438 images from 14 different plants. Figure 2 randomly selects and displays several different categories of diseased plant leaves for illustration.

Figure 2: Dataset Exhibition

2.2. Deep learning models

This study investigates the impact of four different transfer learning models on plant disease detection using deep neural networks on the same dataset, Plant Village. Transfer learning is a commonly used machine learning technique in contemporary times, where knowledge gained from training in one domain is applied to another new domain for further training. Generally speaking, in deep learning, the first few layers are trained to define task-specific features, while the final layers can be removed or new layers added for retraining on the new target task, thus saving significant time and potentially improving accuracy in the new task. This study primarily compares the performance of the Efficientnet_b3a architecture with state-of-the-art CNNs such as Resnet, Google Net, and Inception_V3. Table 1 presents the advantages and disadvantages of each of the four models. In this study, four metrics will be considered to evaluate the results, including precision, recall, and other metrics.

1) Efficientnet_b3a

Since 2014, with the continuous expansion of the ImageNet dataset, the models used have become increasingly complex, leading to significant improvements in accuracy. However, many of these models still have considerable room for improvement in computational efficiency. EfficientNet, as one of the most advanced models today, can be viewed as a family of CNN models. EfficientNet introduces a parameter called compound coefficient, phi, which controls the overall scale of the network. By adjusting the value of phi, various sizes of networks can be constructed under different computational resource constraints to meet the needs of different tasks. Even when the model size increases, the number of parameters does not increase significantly, but the accuracy can be significantly improved. Additionally, EfficientNet differs from other CNN models in that it uses a new activation function called Swish instead of the Rectified Linear Unit (ReLU) activation function. A schematic diagram of the Efficientnet_b3a model is shown in Figure 3.

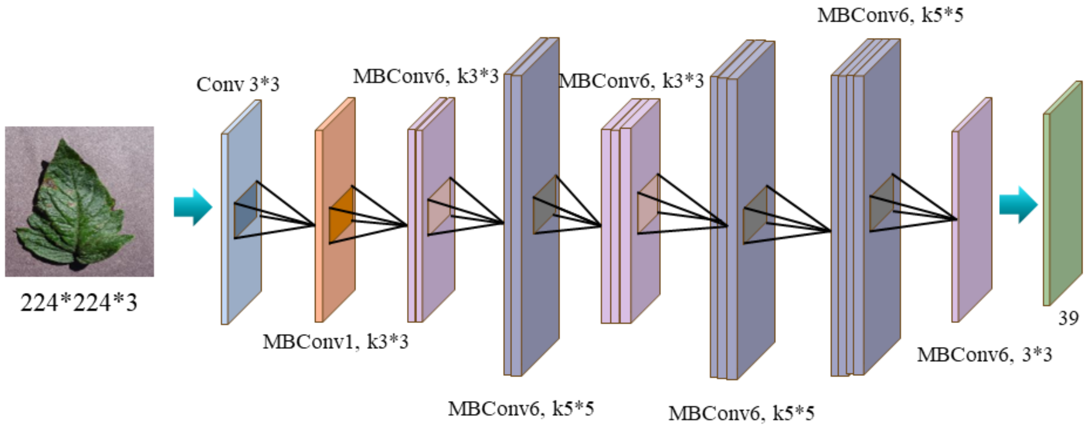

Figure 3: Schematic Diagram of the Efficientnet_b3a Model

2) Resnet50

The Resnet model was proposed by researchers at Microsoft Research in 2015. After the Alex Net model emerged as a top performer in the ImageNet competition in 2012, almost all other network model architectures aimed to improve accuracy by using more layers, but they faced a crucial problem: vanishing/exploding gradients. Resnet addressed this challenge by introducing residual networks. The schematic diagram of the Resnet50 model is shown in Figure 4.

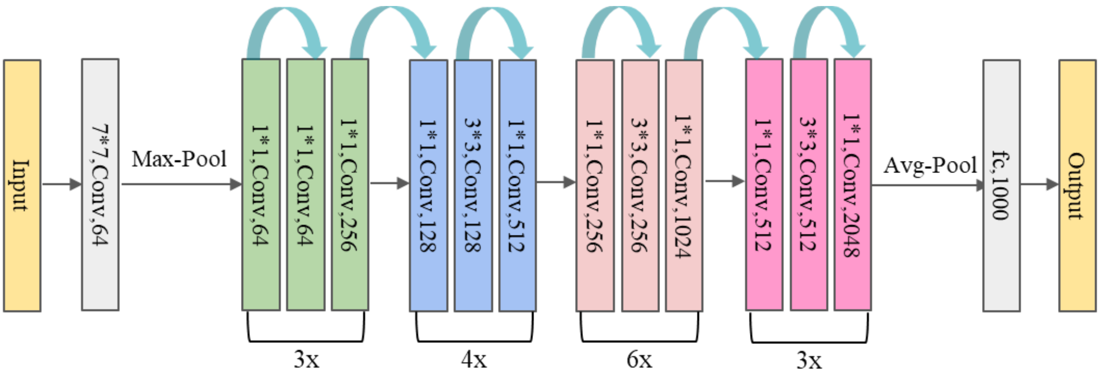

Figure 4: Schematic Diagram of the Resnet50 Model

3) Google Net

The GoogleNet model, as a deep learning model, was introduced after the ResNet model in the ImageNet Large Scale Visual Recognition Challenge 2014. The GoogleNet model consists of 22 layers and is designed as an exceptional image classification model. The initial module of this model is called Inception, and there are nine such initial modules in total. Each initial module possesses convolutional and max pooling layers of different sizes. In these modules, the initial modules process the input through multiple convolutional filters and then concatenate the results together, enabling the network to extract features from multiple levels and cover a wide range of areas in the image. A schematic diagram of the GoogleNet model is shown in Figure 5.

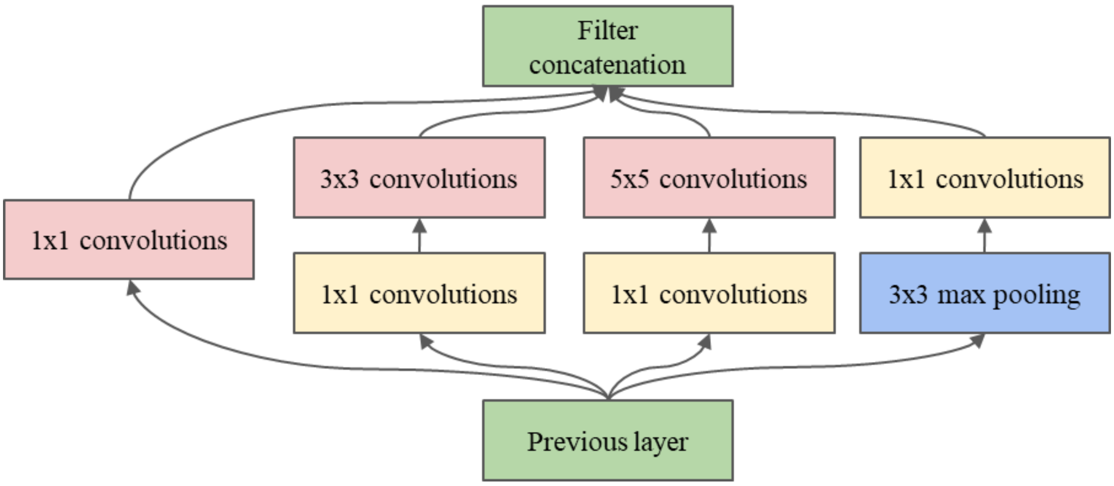

Figure 5: Schematic Diagram of the GoogleNet Model

4) Inception_V3

Inception_V3 is another evolutionary deep-learning architecture. Developed by researchers at Google in 2016, Inception_V3 builds upon the Inception_V1 created by Szegedy. Compared to its previous versions, a key innovation of Inception_V3 lies in the introduction of factorization, a concept used to reduce the number of connections and parameters while maintaining network efficiency. Additionally, a distinguishing feature of Inception_V3 compared to the other three architectures studied in this research is its input layer, which accepts images with a resolution of 299×299 pixels, whereas the other three architectures use a resolution of 224×224. The schematic diagram of the Inception_V3 model is shown in Figure 6.

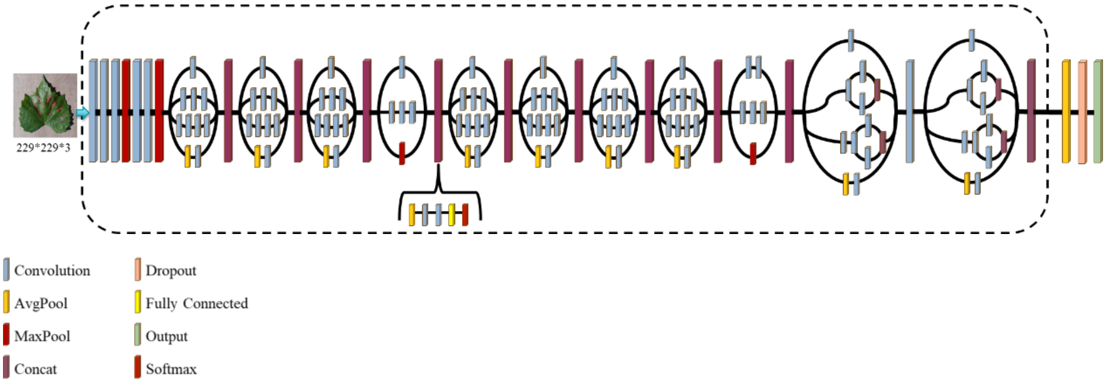

Figure 6: Schematic Diagram of the Inception_V3 Model

Table 1 Comparison of neural networks in common use table

Models	Advantage	Shortcoming
Efficientnet_b3a	High performance, lightweight, portability	Numerous versions
Resnet50	Residual thought, clear structure, and strong expansibility	Long training time
Google Net	Modular structure, which solves the problem of gradient disappearance caused by increasing network depth	Complex model
Inception_V3	Enhanced model expression capability to handle richer spatial features	Complex architecture

3. Experiments

3.1. Experimental Setup

All training experiments in this study were compiled based on a 64-bit Windows 10-22621.3155 system with a GPU platform. The hardware configuration of the operating system includes a CPU of 13th Gen Intel(R) Core(TM) i7-13700F @ 2.10Hz, a GPU of NVIDIA GeForce GTX 1650 with driver version 31.0.15.3623, 12GB of memory, and 16GB of RAM. The software utilizes the PyCharm platform to configure an Anaconda virtual environment. The specific software packages and their versions are cuda12.1, torch 2.2.1+cu121, torchaudio2.2.1+cu121, and torchvision0.17.1+cu121.

3.2. Training

In this study, the plant disease detection system is divided into three stages: training, validation, and testing. The dataset is categorized into three parts, with 60% as the training set, 20% as the validation set, and the remaining 20% as the testing set. The dataset will be randomly divided according to these proportions. The training set and validation set will be used exclusively for model training and fitting, while the testing set will be utilized to assess the generalization capability of the trained model. Table 2 presents the specific details of the dataset partitioning based on the aforementioned proportions.

Table 2: Division of Images into Training Set, Validation Set, and Test Set

	Total	Train	Validation	Test
Original dataset	55438	33289	11085	11064

In this study, to improve the accuracy and training speed of the models, various adjustments were made to different architectures. Specifically, the fully connected layers in the last layer of Efficientnet_b3a, Resnet50, GoogleNet, and Inception_V3 models were replaced with new linear layers, and the number of output features was modified to 39, consistent with the number of data categories, as shown in Figure 7. Additionally, the learning rate was set to 1e-3, the Adam optimizer was used, and the learning rate decay was set to 1e-4. Table 3 presents the specific parameter settings for each model. Furthermore, based on the default sizes acceptable to different models, the image input for Inception_V3 was set to 299×299 pixels, while the image inputs for the other three architectures were set to 224×224 pixels. Table 3 summarizes the default image resolutions and the number of parameters defined for the deep learning models. Table 4 showcases the default input images and total parameters for the deep learning models. Figure 8 displays specific data and comparisons during the training process of the four models.

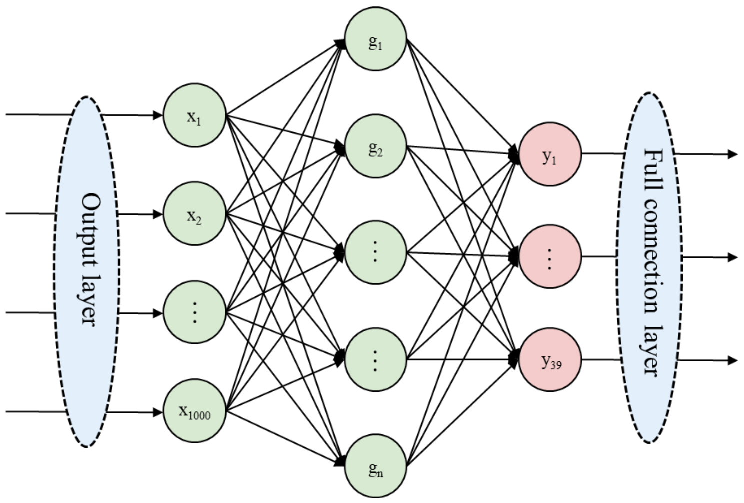

Figure 7: Diagram of Model Structure Adjustment

Table 3: Model Parameter Settings

Parameter	Value
img_size	224
lr	1e-3
batch size	4
epochs	20
weight decay	1e-4

Table 4: Default Input Images and Total Parameters of the Deep Learning Model

Model name	Input size	Number of total parameters
Efficientnet_b3a	224×224	1,624,296
Resnet50	224×224	25,557,032
Google Net	224×224	6,624,904
Inception_V3	299×299	23,834,568

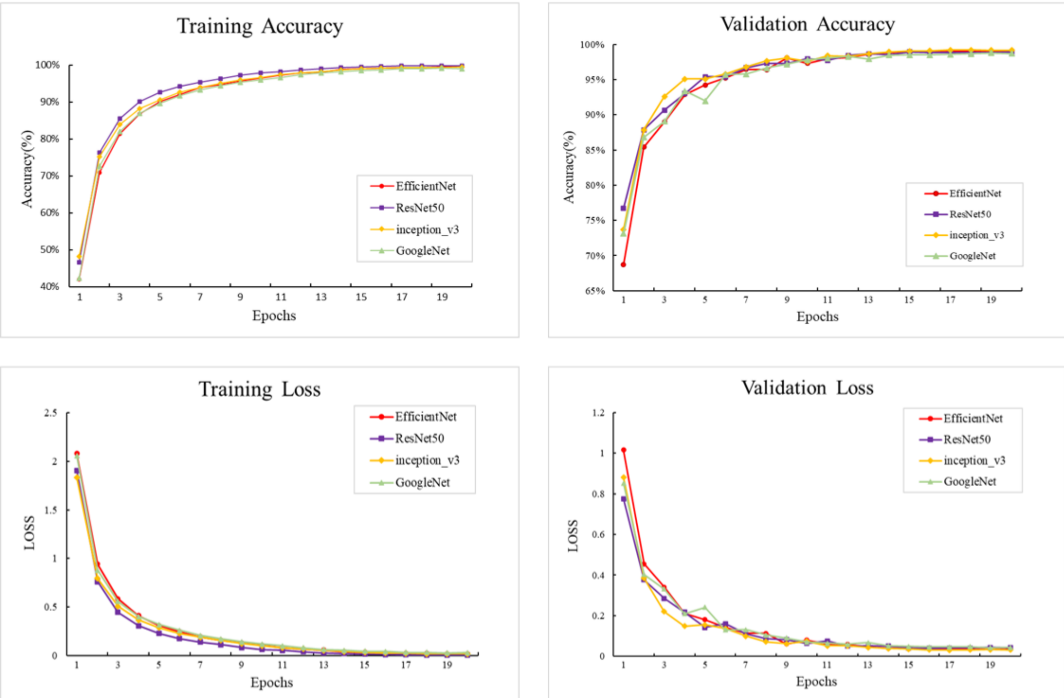

Figure 8: Comparison Data

3.3. Performance Index

Different training models in this study were evaluated using four parameters: Accuracy, Recall, Precision, and F1 score. In addition, heatmaps were employed to demonstrate the testing effectiveness of each model, leading to reasonable conclusions.

A. Accuracy

Accuracy is defined as the ratio of the number of correctly predicted images to the total number of images predicted by the model.

$$accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

B. Recall

Recall, also known as the true positive rate, is the proportion of correctly classified positive samples among all positive samples.

$$recall = \frac{TP}{TP + FN}$$

C. Precision

Precision, also known as the positive predictive value, refers to the proportion of truly positive samples among all samples predicted as positive by the model.

$$precison = \frac{TP}{TP + FP}$$

D. F1-Score

The F1 Score is a metric used in statistics to measure the accuracy of binary classification (or multi-task binary classification) models. It takes into account both the precision and recall of the classification model. The F1 Score can be viewed as a weighted average of the model's precision and recall, with a maximum value of 1 and a minimum value of 0. A higher F1 Score indicates a better performing model.

$$F1\,\mathrm{score} = 2 * \left[\left(precison * recall\right) / \left(precision + recall\right)\right]$$

3.4. Results and Discussion

After training the original divided training set, this study tested 11064 images to obtain the classification performance of the four models as shown in Table 5. Among them, the Efficientnet_b3a model achieved a recall rate of 98.84%, an F1 score of 98.84%, and an accuracy of 99.13%. Compared to the other three models, this network model exhibited better prediction performance.

Table 5: Classification Performance of Different Models

Model	Recall	F1-score	Accuracy
Efficientnet_b3a	98.84%	98.84%	99.13%
Resnet50	98.94%	98.93%	99.19%
Google Net	98.64%	98.64%	98.96%
InceptionV3	98.99%	98.99%	99.24%

Additionally, in this study, 12 categories of crop images were randomly selected from the Plant Village dataset (Table 6 shows the specific variable substitutions) for validation. The four models mentioned in this article were used for testing, and the corresponding heatmaps were generated as shown in Figure 9 (from left to right, top to bottom: Efficientnet_b3a, Resnet50, GoogleNet, and Inception V3). From the heatmaps, it can be observed that when using the test set images for plant disease recognition and classification, the Efficientnet_b3 model achieved an average prediction accuracy of 99.1%, indicating excellent recognition performance and reflecting its suitability for detecting plant leaf diseases.

Table 6: Variable Substitution

New variable	Original variable
x1	Apple___Cedar_apple_rust
x2	Corn___Common_rust
x3	Grape___Leaf_blight_(Isariopsis_Leaf_Spot)
x4	Peach___Bacterial_spot
x5	Strawberry___healthy
x6	Tomato___Septoria_leaf_spot

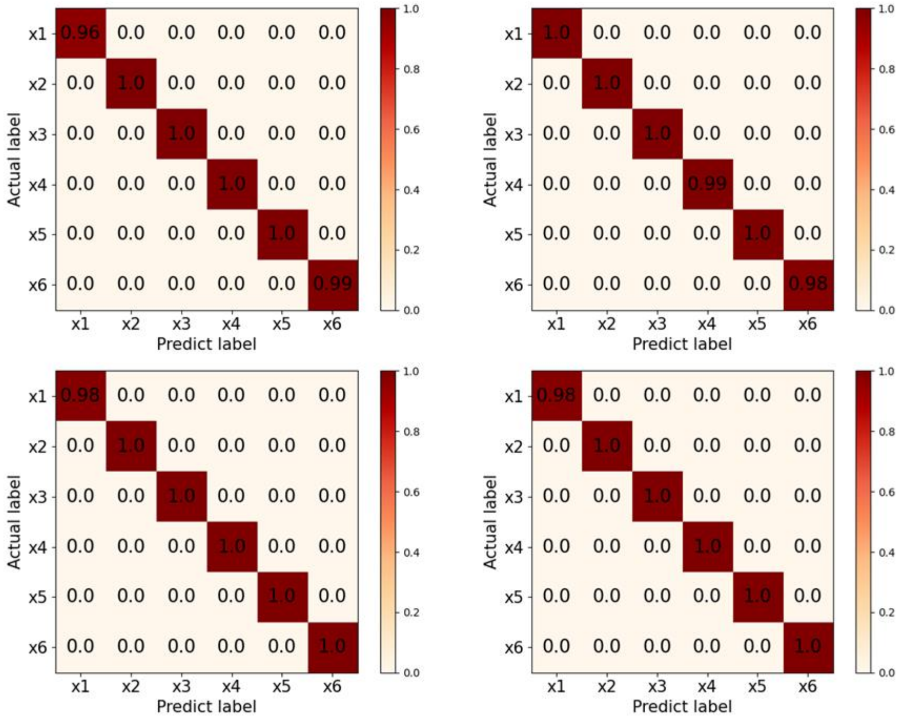

Figure 9: Heatmap Analysis

4. Conclusion

To address the issues of missed detection, low efficiency, and high time consumption in manually identifying diseased plant leaves, we innovatively proposed the use of the Efficientnet_b3a deep learning model for detecting diseased plant leaves. Through experimental validation using the public dataset Plant Village for plant disease leaves, the following main conclusions were drawn: Firstly, the Efficientnet_b3a model used in this study exhibited excellent performance on the dataset, achieving an accuracy rate of up to 99.13%. This accuracy is comparable to that of other traditional networks such as Resnet50, Inception V3, and Google Net, indicating its suitability for plant disease leaf detection. Secondly, the Efficientnet_b3a model used in this study is a lightweight model with good network complexity. While maintaining excellent accuracy, its own model is lightweight, which is very conducive to the subsequent deployment in the hardware system, reducing the deployment cost and computing resource requirements on embedded devices or edge computing platforms, and thus realizing the rapid detection of various plant diseases, helping farmers to minimize the impact of diseases on crop yield and quality, and improving crop production efficiency and economic benefits.

References

[1] Li, Lili, Shujuan Zhang, and Bin Wang. "Plant disease detection and classification by deep learning—a review." IEEE Access 9 (2021): 56683-56698.

[2] Y. Zhou, A. Rangarajan and P. D. Gader, "An Integrated Approach to Registration and Fusion of Hyperspectral and Multispectral Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 5, pp. 3020-3033, May 2020.

[3] J. Choi, K. Yu and Y. Kim, "A New Adaptive Component-Substitution-Based Satellite Image Fusion by Using Partial Replacement," in IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 1, pp. 295-309, Jan. 2011.

[4] Ganaie, Mudasir A., et al. "Ensemble deep learning: A review." Engineering Applications of Artificial Intelligence 115 (2022): 105151.

[5] Jiang, Bo, et al. "Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues." Artificial Intelligence in Agriculture 1 (2019): 1-8.

[6] G. Zhang, S. Zhao, W. Li, Q. Du, Q. Ran et al., "HTD-Net: A deep convolutional neural network for target detection in hyperspectral imagery," Remote Sensing, vol. 12, no. 9, pp. 1–21, 2020.

[7] W. Ma, Q. Guo, W. Yue, W. Zhao, X. Zhang et al., "A novel multi-model decision fusion network for object detection in remote sensing images," Remote Sensing, vol. 11, no. 7, pp. 1–18, 2019.

[8] Bansal, P., Kumar, R., & Kumar, S. (2021). Disease Detection in Apple Leaves Using Deep Convolutional Neural Network. Agriculture, 11(7), 617. doi:10.3390/agriculture11070617

[9] Deepa, R. N and C. Shetty, "A Machine Learning Technique for Identification of Plant Diseases in Leaves," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 481-484, doi: 10.1109/ICICT50816.2021.9358797.

[10] KP, A., & Anitha, J. (2021). Plant disease classification using deep learning. 2021 3rd International Conference on Signal Processing and Communication (ICPSC). doi:10.1109/icspc51351.2021.945169