Proceedings of the 2nd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/74/20240477

Detection and classification of pests and diseases in rice leaf images based on deep learning algorithms

Yiran Cao

School of Software, Xinjiang University, Xinjiang, Urumqi, 830046, China

Corresponding author email: 1792913294@qq.com

Abstract. Rice is one of the important food crops worldwide, but it is susceptible to various pests and diseases during its growth process, leading to yield reduction and even death. Therefore, it is very important to detect and categorize rice pests and diseases in a timely and accurate manner. In this paper, we use the dataset from Kaggle public database, which is divided into two categories of images: "normal leaves" and "infested by pests and diseases", and configure AlexNet, Vgg and MobileNetV2 models based on Pytorch framework to detect and classify rice pests and diseases in a timely and accurate manner. MobileNetV2 model was configured based on Pytorch framework to detect and classify pests and diseases in rice leaf images. Through the changes of the loss and Accuracy curves of AlexNet, Vgg and MobileNetV2 models, we found that all three deep learning models have good training effect and can well predict whether the rice leaf images are infested by pests and diseases. The final results show that the prediction accuracy of all three sets of deep learning models reaches more than 97%, and they can detect pests and diseases on rice leaf images very accurately. This means that we can use these models to quickly and accurately detect and categorize rice pests and diseases, so that timely measures can be taken to prevent the spread of pests and diseases and improve the yield and quality of rice.

Keywords: AlexNet, MobileNetV2, Pests and diseases.

1. Introduction

Rice is one of the important food crops worldwide, but it is susceptible to various pests and diseases during its growth process, leading to yield reduction and even death [1]. Therefore, timely and accurate detection and classification of rice pests and diseases are very important. Traditional rice pest and disease detection methods usually require visual diagnosis by professionals, which is time-consuming and not accurate enough [2,3]. In recent years, with the development of deep learning technology, more and more studies have begun to explore the use of computer vision and deep learning techniques to realize automated rice pest detection and classification [4].

Deep learning techniques can automatically extract image features and classify them by training on a large amount of data. In rice leaf image pest detection and classification, deep learning models can recognize and distinguish different types of pests and diseases by training a large number of rice leaf image datasets with high accuracy and reliability [5].

Automated detection and classification of rice pests and diseases can be realized by using deep learning technology, which can greatly improve the detection efficiency and accuracy, and reduce the cost and human resource requirements [6]. Secondly, timely detection and treatment of rice pest and

DOI: 10.54254/2755-2721/74/20240477

disease problems can ensure rice yield and quality, which is important for ensuring global food security [7]. In this paper, based on the publicly available dataset of rice leaf images, healthy leaf images as well as leaf images infested with pests and diseases were selected, and AleNet, Vgg, and MobileNetV2 models were applied to detect pests and diseases in rice leaf images, respectively, to categorize the images into two classes: healthy and infested with pests and diseases and to compare the prediction of each deep modeling algorithm for the classification of pests and diseases in rice leaf images Effectiveness. This study can provide reference and inspiration for the detection and classification of pests and diseases in other crops.

2. Data presentation and data distribution

The dataset used in this paper is taken from the Kaggle public database and is categorized into two types of images, "normal leaves" and "infested by pests and diseases". The dataset is divided into training set and validation set according to the proportion, in which 1206 images of "normal leaves" and 1304 images of "infested by pests and diseases" are selected as the training set, 134 images of "normal leaves" and 144 images of "infested by pests and diseases" are selected as the training set. 134 images of "normal leaves" and 144 images of "infested by pests and diseases" are selected as the validation set. Figure 1 shows some of the "normal leaf" images and Figure 2 shows some of the "infested" images.

Figure 1. Partial "normal leaf" image.
(Photo credit: Original)

Figure 2. Part of the "infested" image.
(Photo credit: Original)

3. Method

3.1. AlexNet

AlexNet is a milestone in the field of deep learning, it is the winner of the 2012 ImageNet Large Scale Visual Recognition Competition (ILSVRC) and the first convolutional neural network model to apply deep learning to computer vision tasks.The AlexNet model was developed by Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton [8], and its main principles include the following.

DOI: 10.54254/2755-2721/74/20240477

AlexNet utilizes a multilayer convolutional neural network structure. In traditional image classification models, shallow fully connected neural networks or SVMs are usually used for classification. In contrast, AlexNet can automatically learn image features by introducing a multilayer convolutional neural network structure, and avoids the tedious process of manually designing features.AlexNet uses the ReLU activation function. In traditional neural networks, the sigmoid function is usually used as the activation function. However, the sigmoid function has the problem of vanishing gradient in the derivation, which leads to a very slow training process. The ReLU activation function can effectively solve this problem, and can also accelerate the training process.

AlexNet uses data enhancement techniques. Data enhancement techniques can expand the training set and reduce the occurrence of overfitting.AlexNet uses a variety of data enhancement techniques, including random cropping, level flipping, color transformation, etc. AlexNet uses Dropout.AlexNet uses Dropout, which removes neurons randomly during training to minimize overfitting, in both the fully connected and convolutional layers.

AlexNet uses GPU parallel computing technique. When training deep neural networks, it usually requires a lot of computational resources and time. And GPU parallel computing technology can accelerate the training process of the model and can handle larger datasets.

In conclusion, the AlexNet model is of great importance in the field of deep learning, and AlexNet has successfully solved many problems in image classification tasks and achieved very good results.

3.2. Vgg

The VGG model is an important model in the field of deep learning, which was proposed by a team of researchers from the University of Oxford and won the second place in the 2014 ImageNet Large Scale Visual Recognition Competition (ILSVRC) [9].The main principles of the VGG model include the following aspects.

The VGG model uses a very deep convolutional neural network structure. In traditional convolutional neural networks, 5-7 layers of convolutional layers and 2-3 layers of fully connected layers are usually used for feature extraction and classification. The VGG model, on the other hand, employs 16-19 layers of convolutional layers and 3 layers of fully connected layers, allowing the model to learn image features more fully.The VGG model employs a small size convolutional kernel. In traditional convolutional neural networks, a convolutional kernel of 3x3 or 5x5 size is usually used for feature extraction. The VGG model uses a small size (3x3) convolutional kernel and uses multiple small size convolutional kernels stacked instead of one large size (e.g., 7x7) convolutional kernel, which reduces the number of parameters and better captures the local features in the image.

The VGG model uses a pooling layer. The pooling layer reduces the computation and number of parameters by downscaling the feature map output from the convolutional layer. In VGG model, a maximum pooling layer of 2x2 size is usually used for dimensionality reduction and a sliding window of step size 2 is used for scanning.The VGG model uses the ReLU activation function, which speeds up the training process and avoids the gradient vanishing problem. In VGG model, all the convolutional and fully connected layers use ReLU activation function.

The VGG model uses Dropout and Data Enhancement techniques, which can randomly delete some neurons during the training process to reduce the occurrence of overfitting, and Data Enhancement techniques, which can expand the training set and reduce the occurrence of overfitting.

In conclusion, the VGG model is a very classical and important deep learning model. By using deep convolutional neural network structure, small size convolutional kernel, pooling layer, ReLU activation function, Dropout technique and data enhancement technique, the VGG model successfully solves many problems in the image classification task and achieves very good results.

3.3. MobileNetV2

MobileNetV2 is a lightweight model in the field of deep learning, which was proposed by a research team at Google and has received very good results in the 2018 ImageNet Large Scale Visual Recognition Competition (ILSVRC) [10].The main principles of the MobileNetV2 model include the following.

DOI: 10.54254/2755-2721/74/20240477

MobileNetV2 uses a depth-separable convolution structure. Depth separable convolution is a new type of convolutional neural network structure, which divides the standard convolution into two steps: depth convolution and point-by-point convolution. Deep convolution can process the information on each channel separately, thus reducing the amount of computation; point-by-point convolution can interact and integrate the information between different channels, thus preserving the feature information. By adopting a depth-separable convolution structure, MobileNetV2 can significantly reduce the number of parameters and the amount of computation while guaranteeing higher accuracy.

MobileNetV2 adopts a linear bottleneck structure. The linear bottleneck structure is a new type of residual block structure, which replaces the nonlinear activation function in the traditional residual block with a linear transformation and adds a 1x1 convolution layer before the point-by-point convolution. By adopting a linear bottleneck structure, MobileNetV2 can further reduce the amount of computation and the number of parameters, and can maintain high accuracy.

MobileNetV2 adopts an inverted residual structure. The inverted residual structure is a new type of residual block structure, which changes the shortcut connection from input to output to output to input in the traditional residual block, thus avoiding the problem of information loss and degradation. By using the inverted residual structure, MobileNetV2 can better retain feature information and can improve model accuracy.MobileNetV2 uses lightweight design principles. During the model design process, the Google team used some lightweight design principles, such as channel number factorization, grouped convolution, dynamic convolution, and so on. These design principles can further reduce the amount of computation and the number of parameters, and can maintain high accuracy.

MobileNetV2 adopts the SE module, which is a new type of attention mechanism, and the introduction of SE module in convolutional neural networks can effectively improve the feature expression ability and classification accuracy.MobileNetV2 adds an SE module after each depth-separable convolution.

In conclusion, MobileNetV2 is a very lightweight and efficient deep learning model. By adopting the depth-separable convolutional structure, linear bottleneck structure, inverted residual structure, lightweight design principle, and SE module, MobileNetV2 successfully solves the image classification task in scenarios with limited computational resources and achieves very good results. Meanwhile, MobileNetV2 also provides a lot of inspiration and reference for later deep learning models.

4. Experiments and Results

4.1. Experiments

After dividing the dataset, the AleNet, Vgg and MobileNetV2 models are configured based on the Pytorch framework, respectively, and a series of pre-processing such as cropping, rotating and normalizing are performed on the input images (including the training set and the validation set), the epoch is set to 20, and the learning rate is set to 0.0001 after debugging, and the batch size is set to 32, and the records of the The change curve of loss in the training set and the change of Accuracy in the validation set for each model.

4.2. Results

Record and plot the change in the loss curve of the training set and the change in the Accuracy of the validation set of each model, the change in the loss curve of the training set and the change in the Accuracy of the validation set of the AlexNet model are shown in Fig. 3, the change in the loss curve of the training set and the change in the Accuracy of the validation set of the Vgg model are shown in Fig. 4, and the change in the Accuracy of the MobileNetV2 model is shown in Fig. 5. The loss variation curve for the training set of MobileNetV2 model and the Accuracy variation curve for the validation set are shown in Fig. 5.

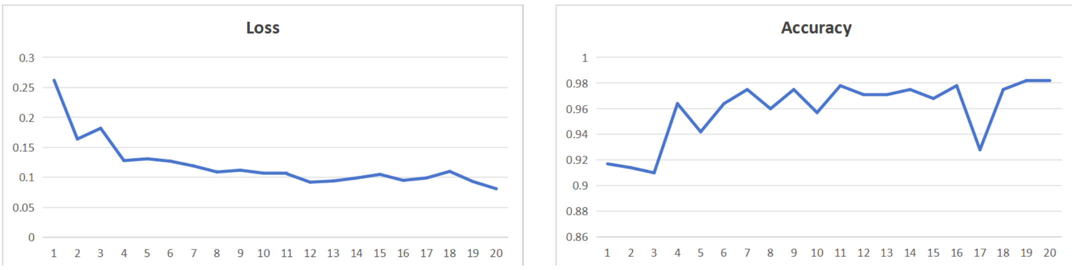

Figure 3. Loss variation curves for the training set and Accuracy variation curves for the validation set of the AlexNet model.

(Photo credit : Original)

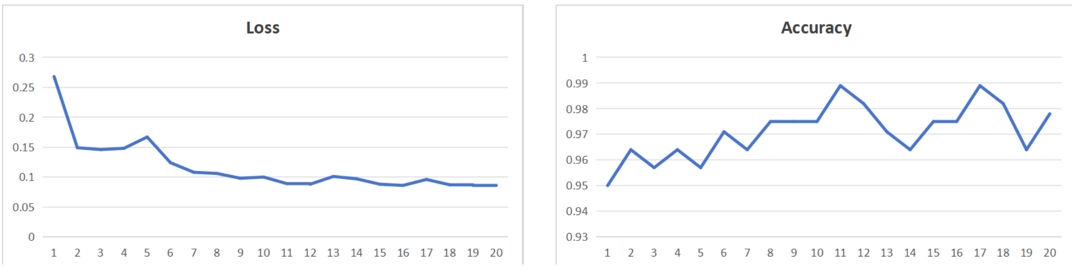

Figure 4. Loss variation curves for the training set and Accuracy variation curves for the validation set of the Vgg model.

(Photo credit : Original)

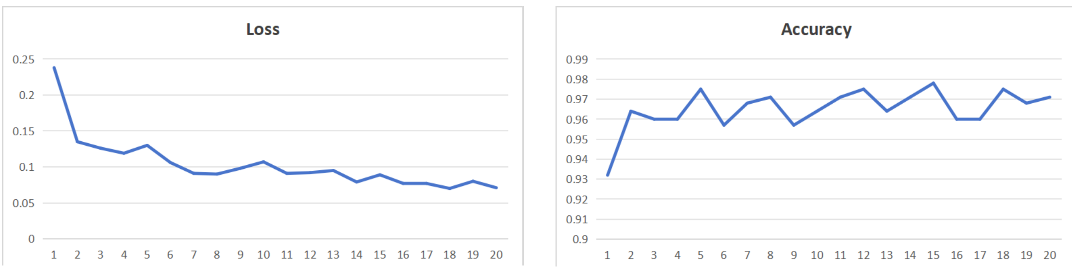

Figure 5. Loss variation curves for the training set and Accuracy variation curves for the validation set of the MobileNetV2 model.

(Photo credit : Original)

From the changes in the loss and Accuracy curves of AlexNet, Vgg and MobileNetV2 models, it can be seen that as the training process proceeds, the loss gradually decreases and tends to converge, and the Accuracy gradually increases and regionally converges, and a few epochs with the best prediction results of the three models are recorded, as shown in Table 1.

Table 1. Model evaluation indicators.

Model	Best epoch	Accuracy(%)
AlexNet	19	98.2
Vgg	11	98.9
MobileNetV2	15	97.8

From the training results, it can be seen that the prediction accuracy of all three sets of deep learning models reaches more than 97%, which can be very accurate in detecting pests and diseases on rice leaf images.

5. Conclusion

Rice is one of the important food crops in the world, but it is susceptible to various pests and diseases during the growth process, leading to yield reduction and even death. Therefore, it is important to detect and categorize rice pests and diseases in a timely and accurate manner. In this paper, we configured AleNet, Vgg and MobileNetV2 models based on Pytorch framework to detect and classify rice pests and diseases in a timely and accurate manner. The MobileNetV2 model was configured based on the Pytorch framework to detect and classify pests and diseases in rice leaf images.

Through the changes of loss and accuracy curves of AlexNet, Vgg and MobileNetV2 models, we found that as the training process proceeds, the loss gradually decreases and tends to converge, and the accuracy gradually increases and regionally converges. This indicates that all three deep learning models have good training results and can well predict whether rice leaf images are infested with pests or diseases. The final results show that the prediction accuracy of all three groups of deep learning models reaches more than 97%, which can detect pests and diseases on rice leaf images very accurately. This means that we can use these models to quickly and accurately detect and categorize rice pests and diseases, so that timely measures can be taken to prevent the spread of pests and diseases and improve the yield and quality of rice.

In conclusion, the research results in this paper provide strong technical support for the rice farming industry. In the future, we will continue to optimize the model algorithms, further expand the dataset size, and further improve the accuracy and efficiency of rice pest and disease detection and classification.

Acknowledge

This project was supported by the Xinjiang University Project (202310755066).

References

[1] Alpyssov A ,Uzakkyzy N ,Talgatbek A , et al.Assessment of plant disease detection by deep learning[J].Eastern-European Journal of Enterprise Technologies,2023,1(2):41-48.
[2] Ruyi Z ,Fanhuai S .A novel strategy for pest disease detection of;Brassica chinensis;based on UAV imagery and deep learning[J].International Journal of Remote Sensing,2022, 43(19-24):7083-7103.
[3] Liang D ,Liu W ,Zhao Y .Optimal Models for Plant Disease and Pest Detection Using UAV Image[J].Nature Environment and Pollution Technology,2022,21(4):1609-1617.
[4] Dawei L ,Foysal A ,Nailong W , et al.YOLO-JD: A Deep Learning Network for Jute Diseases and Pests Detection from Images[J].Plants,2022,11(7):937-937.
[5] Fan T .Research on Edge Detection of Agricultural Pest and Disease Leaf Image Based on LVQ Neural Network[J].Recent Advances in Computer Science and Communications, 2021,14(6):1903-1911.
[6] Mahalakshmi D S ,Vijayalakshmi K .Agro Suraksha: pest and disease detection for corn field using image analysis[J].Journal of Ambient Intelligence and Humanized Computing,2020, 12(prepublish):1-15.
[7] Zhang K ,Sun W ,Ba Y , et al.Transformer Fault Diagnosis Method Based on SCA-VMD and Improved GoogLeNet[J].Applied Sciences,2024,14(2):
[8] Shi L ,Hai B ,Kuang Z , et al.ResnetAge: A Resnet-Based DNA Methylation Age Prediction Method[J].Bioengineering,2023,11(1):
[9] Abbass B A M ,Ban Y .MobileNet-Based Architecture for Distracted Human Driver Detection of Autonomous Cars[J].Electronics,2024,13(2):
[10] Zhiyao Y ,Xu Y ,Heng Z , et al.An Android Malware Detection Method Using Multi-Feature and MobileNet[J].Journal of Circuits, Systems and Computers,2023,32(17):