ECA-ConvNeXt: A Rice Leaf Disease Identification Model Based on ConvNeXt

Xiaoqi Wang, Yaojun Wang[*], Jingbo Zhao[*], Jing Niu

College of Information and Electrical Engineering, China Agricultural University

Beijing, China

{wangyaojun, zhao.jingbo}@cau.edu.cn

Abstract

As an essential food crop, rice is often infested with diseases that can cause significant yield losses and seriously damage economic income and food health. Early identification and control of rice diseases is an effective way to alleviate these problems. However, manual identification and diagnosis of rice leaf diseases requires experienced specialists and is time-consuming. In our study, we propose the ECA-ConvNeXt model, based on the ConvNeXt network, which can identify six categories of typical rice leaf diseases and healthy rice leaves. We also established a rice leaf disease identification dataset that contains images of healthy and diseased rice leaves with complex backgrounds and their disease category labels. In the proposed ECA-ConvNeXt model, we incorporated the ECA (Efficient Channel Attention) module, which improved the feature extraction performance using only a few parameters. Transfer learning was applied to load pre-training weights and fine-tuning was used to reduce training costs and improve the model performance. We tested the performance of ECA-ConvNeXt on the rice leaf disease identification dataset. Experimental results show that the proposed model achieved an accuracy of 94.82%, a precision of 94.47%, a recall rate of 94.31%, and an F1-Score of 94.33% on the rice leaf disease identification dataset. These results suggest that the proposed network effectively identifies rice leaf diseases.

1. Introduction

Rice is one of the most important human food crops, with abundant edible, economic, and medicinal values. According to statistics, half of the world's population consumes rice, and the total rice production ranks third in the world for food crop production, following corn and wheat [1]. However, all major rice production areas worldwide have suffered from rice diseases for a long time [2]. The spread of diseases can significantly reduce rice yield and quality and threaten the food supply. Therefore, reliable and rapid diagnosis of rice diseases and preventive

and control measures are crucial for disease control. Artificial Intelligence (AI) technology and targeted treatments offer new ways to identify rice diseases to ensure healthy rice growth accurately and are thus worth investigating.

The emergence and development of computer vision and deep learning techniques have opened new directions for precision and smart agriculture. They also provide a new solution to the problem of identification and diagnosis of plant diseases. In the past few years, with the advances in digital camera technology and the improvement of computing power, image recognition methods have demonstrated superior performance in various applications and have been widely used in agriculture [3]. Currently, the identification of crop diseases is mainly accomplished by farmers and agricultural professionals. Manual identification methods are mainly derived from farmers' experience, books, search engines, and the guidance of professionals [4]. Nevertheless, these approaches are time-consuming, laborious, economically costly, and have low accuracy. Therefore, with the rapid development of deep learning methods in image recognition, effective and accurate identification of crop diseases of rice and other crops by vision-based methods has become a popular research topic in crop protection in recent years.

However, the current application of various deep learning models in the field of rice disease identification has yet to achieve the goal of accurately identifying rice leaf diseases. The main research challenge is to identify multiple categories of rice disease with high accuracy. In addition, publicly available rice disease identification datasets are rare and there is limited data for model training and testing. The inadequacy of training samples can make it difficult to train an accurate model. Moreover, images taken in the rice field often suffer from non-uniform environmental backgrounds and multiple leaves can be taken in a single image. These pose challenges to leaf disease identification tasks. When taking images, lighting conditions, camera hardware, and parameter settings of cameras lead to differences in clarity, brightness, and sharpness of captured images. These variations in image quality can make it difficult for deep-learning models to

---

*Corresponding author.

identify and learn features effectively. Furthermore, several diseases may be present on a single rice leaf, which can cause further complications.

Our study proposes a new convolutional neural network (CNN) model for rice leaf disease identification to tackle these research difficulties and challenges. We also tested the model on the rice leaf disease identification dataset to evaluate its performance. The main contributions of our study are:

(1) We proposed a new CNN model, called ECA-ConvNeXt, for rice leaf disease identification. The proposed model can improve diseased rice leaves' classification and identification accuracy in complex backgrounds.

(2) We constructed a dataset using public datasets combined with manual annotation. Seven types of rice leaves were included in the dataset. These are Healthy, Hispa, Brown Spot, Leaf Blast, Bacterial Blight, Bacterial Leaf Streak, and Sheath Blight.

The rest of the article is organized as follows: Section II describes the applications related to CNNs and transfer learning methods in rice disease identification. Section III provides a detailed description of the dataset acquisition and building process. Section IV presents the network structure of ECA-ConvNeXt, including the ECA module and fine-tuning method. Section V presents the performance evaluation of the ECA-ConvNeXt model and the comparison with the state-of-the-art (SOTA) models. Section VI discusses the limitations of our current study and suggests future research. Section VII summarizes our study.

2. Related Work

2.1 Convolutional Neural Networks on Crop Disease Identification

Deep learning methods based on CNNs have been widely used in crop disease identification in recent years. Gupta *et al.* [5] introduced a CNN-based deep learning classification model using the ADAM optimizer to train their model and test the optimal input size of images. Their experimental results showed that when the image input size was 512×512 pixels, the classification accuracy of four categories of diseased rice leaves reached 92.83%. Deng *et al.* [6] selected DenseNet-121, SE-ResNet-50, and ResNeSt-50 as sub-models to construct a recognition model by integrated learning, which minimized the confusion between different types of diseased leaf images. The accuracy of the integrated model for classifying six different rice disease images was up to 91%. Haridasan *et al.* [7] used image segmentation techniques to determine the diseased parts of rice leaves, incorporating a support vector machine (SVM) and a CNN. The experimental results indicated that the model's accuracy for classifying

five types of diseased rice leaves reached 91.45%, alleviating the disease control problem in Indian rice fields. These studies show that CNNs have been widely applied to rice disease identification and have yielded many successes. However, there is still room to improve identification accuracy and increase the number of categories for disease identification.

2.2 Transfer Learning for Rice Disease Identification

Building a deep learning model for rice disease classification and identification requires enormous time and computational resources. Using transfer learning methods to obtain prior knowledge from trained models and transfer them to the training of rice disease identification models can improve the training speed and cope with the problem of limited diseased rice leaf images. Narmadha *et al.* [8] proposed a new deep learning-based rice disease diagnosis model DenseNet169-MLP. Their method utilized DenseNet-169 for transfer learning and replaced the fully connected layer of the DenseNet model with an MLP (Multi-Layer Perceptron) to achieve rice disease classification using the extracted features. The experimental results showed that the classification accuracy of DenseNet169-MLP for three different categories of rice diseases was 97.68%. Chen *et al.* [3] proposed a deep learning framework DENS-INCEP for rice disease identification. The proposed method combined pre-trained DenseNet on ImageNet with the Inception module for transfer learning and introduced the focal loss function to enhance the learning capability of tiny lesion features. The experimental results showed that the accuracy of DENS-INCEP for identifying three different categories of rice diseases in the public dataset UCI [9] was 94.07%. Latif *et al.* [10] proposed a deep convolutional neural network (DCNN) based transfer learning method using VGG19 to extract features and perform feature downscaling to complete the classification of rice diseases with fine-tuning. Their experimental results indicated that this model could accurately detect and diagnose six different categories of rice leaf diseases with the highest average accuracy of 96.08%. These studies show that applying transfer learning methods to rice disease identification models effectively improves accuracy.

Given that there is still room to boost the performance of rice disease identification, our study proposes the ECA-ConvNeXt rice leaf disease identification model based on the latest convolutional neural network ConvNeXt [11]. Our results show that the ECA-ConvNeXt model can complete the accurate identification of rice leaf diseases in seven categories with only a slight increase in model parameters and computational costs.

3. Dataset Collection and Construction

To build a practical and extensive rice leaf disease identification model, we investigated a variety of rice diseases and selected six categories of typical and common rice leaf diseases and healthy rice leaves to build the identification dataset. The dataset contains 4523 raw images and their disease category labels. This section introduces the process of data annotation and data augmentation in detail.

3.1 Image Dataset Collection

There are few publicly available rice leaf disease datasets on the Internet. The experimental data for this study are mainly collected from the Internet, public dataset Rice Leaf Disease Image Samples [12], and relevant rice leaf disease images and labels provided by Kaggle (https://www.kaggle.com). The dataset contains a variety of rice leaves and the images in the dataset were taken under different lighting conditions and at different camera angles. The central parts of the images are the diseased leaves, and the background parts of most images are the natural environment backgrounds with soil and ground. The dataset mainly consists of images with only one type of rice leaf diseases that appear on rice leaves. The collected data were further processed, and our research team members removed duplicated images and labels from different sources.

The dataset contains six categories of rice leaf diseases and healthy rice leaves. Each leaf type has different morphological characteristics, as shown in Figure 1. Healthy leaves are bright, dark, and green, with smooth leaf edges and no spots in the leaves. Hispa leaves have white spots or white longitudinal streaks. In severe cases, the whole rice leaves wither, affecting the fruiting of the spike or even resulting in no fruits. Brown Spot appear on leaves as small oval spots, and the center of the spots is brown. Near the edge, there are different shades of yellow halo. When severe, the halos can be fused into large irregular spots. Leaf Blast appear on leaves as elliptical spots; the spots' edges appear brown, and the middle is gray-white. In humid environments, a gray-green mold layer may appear. Bacterial Blight appears on leaves as yellow-green or dark-green spots in the early stage. Then, along the leaf veins, they quickly expand into stripes. Bacterial Leaf Streak appears on leaves as small and transparent dark brown spots. With the spread of this disease, the spot area gradually becomes more visible. Sheath Blight appears on leaves as dark green spots, and they later expand to oval shapes. Often multiple spots are fused into a large pattern. In dry conditions, the edge of the spot is brown, and the center appears straw yellow to gray.

The exact construction process of the rice leaf disease identification dataset is as follows:

(a) Healthy   (b) Hispa   (c) Brown Spot

(d) Leaf Blast   (e) Bacterial Blight   (f) Bacterial Leaf Streak   (g) Sheath Blight
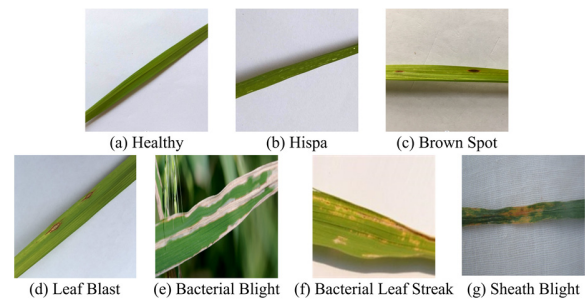
Figure 1: The dataset contains seven types of leaves, including healthy rice leaves and six types of diseased leaves. (a) Healthy Leaf is bright green without diseased spots. (b) Hispa appears on leaves as white spot-like stripes. (c) Brown Spot appears on leaves as a central brownish spot surrounded by a yellow halo near the margin. (d) Leaf Blast appears on leaves as a brown border and grayish-white and elliptical spots in the center. (e) Bacterial Blight appears on leaves as white spots, expands along the leaf, and forms stripes. (f) Bacterial Leaf Streak appears on leaves as a dark yellow water-stained spot. (g) Sheath Blight appears on leaves as brown edges and a yellow straw center.
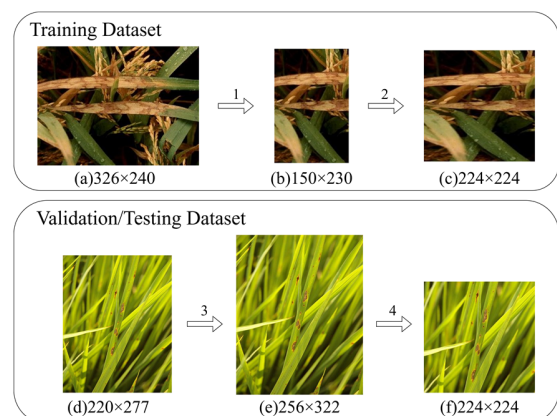

Figure 2: Image pre-processing results. (a) An original image in the training dataset with a size of 326×240. (b) The image is randomly cropped to a size of 150×230. (c) The image is scaled to the fixed size of 224×224. (d) An original image in the validation and testing dataset with a size of 220×277. (e) The image is scaled to 256 pixels on the short edge with a size of 256×322. (f) The image is cropped in the center to the fixed size of 224×224.

(1) Dataset preparation and annotation: we combined the data from the Internet and other publicly available datasets for data aggregation. Under the guidance of plant protection experts, the researchers annotated and checked the images and labels to avoid labeling errors and to remove duplicated images.

(2) Dataset splitting: using random seeds, we split the image samples of each category into a training dataset, a validation dataset, and a testing dataset in a ratio of 6:2:2 to form the experimental dataset.
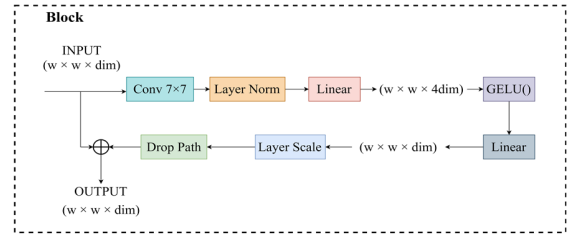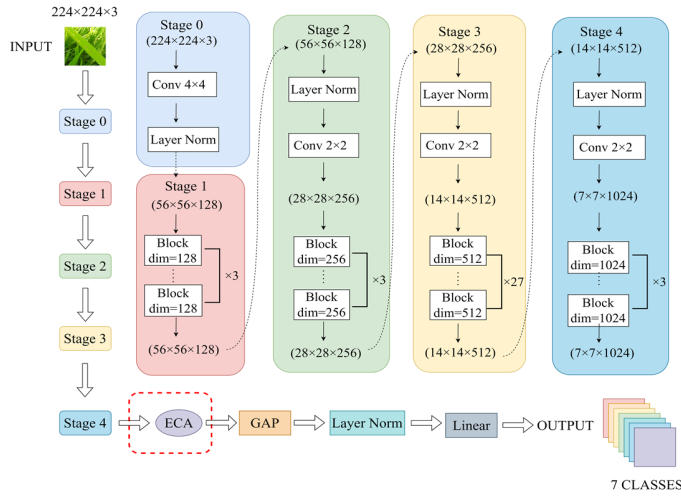
Figure 3: The architecture of ECA-ConvNeXt. The ConvNeXt network is divided into five stages. Each stage is in a different color. The ECA attention modules are added to follow all convolutional layers. The final output is processed by a global average pooling layer and a fully connected layer. The output of the network is a seven-dimensional vector that denotes the predicted probabilities of seven types.

(3) Data augmentation: after the dataset was split, the images and their labels in the training dataset, validation dataset, and testing dataset were augmented in different ways to balance the dataset based on the existing distribution of the images corresponding to their types.

3.2 Data Augmentation and Pre-processing

Data augmentation is an effective method to expand the dataset size, and increasing the dataset size is beneficial to improve the generality of a trained model. We performed data augmentation operations, including "Double flip", "Rotate 120", "Rotate 60", "Horizontal flip", and "Vertical flip" by OpenCV on the dataset based on the number of leaves from different categories to complete the construction of the rice leaf disease identification dataset. The number of images in the training, validation, and testing dataset for different types of leaves is shown in Table 1.

To improve the generality of the rice leaf disease identification model, we converted images of arbitrary sizes into a fixed size of 224×224 as the model input. Specifically, the images in the training dataset were cropped randomly to different sizes and aspect ratios and then scaled to a fixed size of 224×224. The images in the validation and testing dataset were scaled to 256 pixels on their short edge and then cropped in the center to have images of a fixed size of 224×224. The details in the pre-processing can be seen in Figure 2. The last step is the normalization transform, which normalizes the tensor image by the mean and the standard deviation from the ImageNet dataset. The calculation is as in equation (1):

$$output[channel] = \frac{input[channel] - mean[channel]}{std[channel]} \quad (1)$$

where std[channel] denotes the standard deviation of the channel. The values of the mean and the standard deviation are [0.485, 0.456, 0.406] and [0.229, 0.224, 0.225]. The normalization transform can alleviate the gradient disappearance phenomenon during training, improve the learning speed and reduce the dependency of model training on initialization.

4. Methods

In this section, we present the ECA-ConvNeXt in detail. The proposed ECA-ConvNeXt is a CNN architecture for identifying six types of rice leaf diseases and healthy rice leaves based on the ConvNeXt network. We used the ECA attention module, which makes an appropriate modification to the intersection part of the convolutional layers and fully connected layers, to improve the model's feature extraction capability. We also adopted the fine-tuning approach of transfer learning to improve the training efficiency and reduce the time cost. We next introduce the modules of the proposed architecture, as shown in Figure 3.

4.1 Overview of the ECA-ConvNeXt Network

The architecture of ECA-ConvNeXt can be divided into five stages, and the size of the baseline model determines the number of ConvNeXt blocks in each stage. The block composition of the rice leaf disease identification model is (3,3,27,3), which is called ECA-ConvNeXt_Base, in our study. We added the ECA module between the convolutional layers and the global average pooling layer (in Figure 5) to improve the channel attention mechanism of the model, further extract the feature information of the images, and adjust the weight settings of the feature and non-feature channels. Model training adopts the transfer

Name	Training	Validation	Testing
Healthy	864	95	96
Hispa	729	80	81
Brown Spot	834	92	93
Leaf Blast	699	77	78
Bacterial Blight	261	28	29
Bacterial Leaf Streak	168	27	27
Sheath Blight	126	18	21

Table 1: Distribution of rice leaf disease images in the dataset
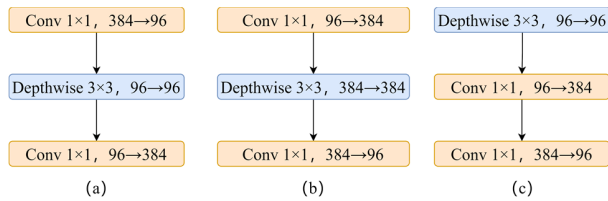

Figure 4: The inverted bottleneck of ConvNeXt and the associated network architecture. (a) In the bottleneck layer of ResNeXt, the depth-wise separable convolution is in the middle layer, and the input and output values are larger than the intermediate values. (b) Improved inverted bottleneck layer, with the depth-wise separable convolution still in the middle layer, and the input and output values are smaller than the intermediate values. (c) The inverted bottleneck of ConvNeXt, where the depth-wise separable convolution is moved up, and the input and output values are still smaller than the intermediate values.

learning method, fine-tuning (in Figure 6). The pre-training weights, completed by ConvNeXt_Base training on ImageNet-1K, are set as the initialization weights of the rice leaf disease identification model to reduce the training costs and improve the model performance.

The main framework of ConvNeXt is based on ResNet50/ResNet200, which is inspired by the Swin Transformer [13]. Liu et al. [11] tested the limits of the pure ConvNets and proposed several vital components to construct and improve the ConvNeXt network. The network was also trained and evaluated on ImageNet-1K to complete the network structure of ConvNeXt. ConvNeXt adjusts the ratio of blocks of each stage to 1:1:3:1 and finally obtains the block ratio as (3,3,9,3), which increases the number of parameters of the model. For larger scale models, the number of blocks is further boosted to (3,3,27,3). ConvNeXt also introduces the idea of group convolution, replacing 3×3 convolution with 3×3 group convolution.

Meanwhile, ConvNeXt learns from ResNeXt [14] and adds an inverted bottleneck to the network structure. To accommodate larger kernel sizes, the depth-wise convolution of the inverted bottleneck layer is shifted up by one layer, and the detailed structure is shown in Figure 4. Finally, the detailed modules of the network, such as activation functions and batch normalization, are further optimized to form the ConvNeXt network structure.

4.2 ECA Module

The images containing diseased leaves are input to the rice leaf disease identification model. Most distinctions in appearances between different types of diseased rice leaves are from the diseased spot areas. Therefore, we added a channel attention mechanism following the convolutional layers to improve the model's identification performance. The attention mechanism can target particular parts of the data and assign different weights to different input components to filter out the significant parts among a large amount of information.

The classical channel attention module is SENet, whose critical operations are the squeeze and the excitation operations. It works by obtaining the importance of each channel of the feature map through the SE module and then using it as a criterion to assign weights to each feature so that the neural network can focus on the feature channels [15]. The core idea is to automatically learn feature weights through fully connected networks and data losses to increase the effective feature channel weights.

Wang et al. [16] argue that the dimensionality reduction operation in the SE module can negatively affect the channel attention mechanism, and capturing the dependencies among all channels is also ineffective. Thus, the ECA module uses 1×1 convolutional layers directly following the global average pooling layer and removes the fully connected layers to avoid dimensionality reduction. The ECA module effectively captures cross-channel interactions and involves only a few parameters to achieve good results. The structural changes of the ECA module to the SE module are shown in Figure 5. An adaptive equation determines the size of the convolution kernel as in (2):

$$k = \left| \frac{log_2\, C}{\gamma} + \frac{b}{\gamma} \right| \quad (2)$$

where k denotes the kernel size and C denotes the channel dimension. γ and b are set to 2 and 1, respectively.

4.3 Fine-tuning

To improve the effectiveness and performance of the rice leaf disease identification model, fine-tuning was used in the training process of the ECA-ConvNeXt. Sharing the relevant model parameters that have been trained on large-scale datasets to the current model can improve learning efficiency and reduce training costs. The fine-tuning method facilitates the learning of parameters by freezing the convolutional layers of the pre-trained model that are used to extract generic features to update the remaining convolutional layers and fully connected layers [17]. Its
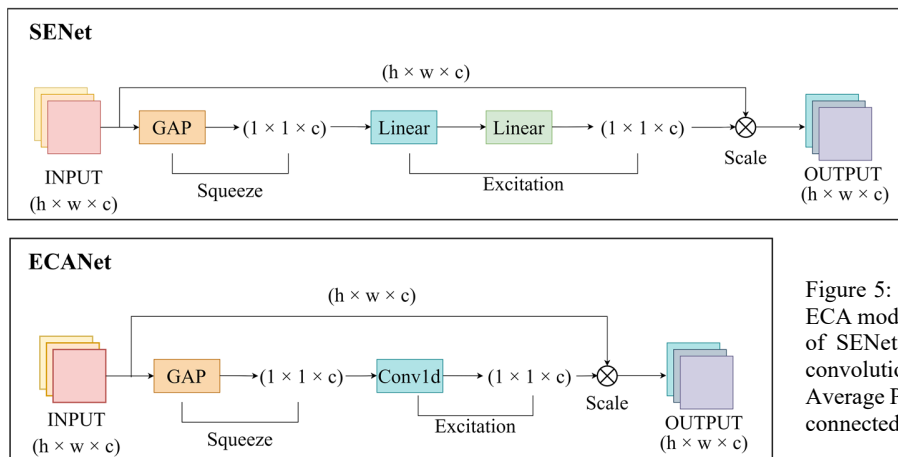
SENet

ECANet

Figure 5: Structure of the SE module and the ECA module. ECANet is an improved version of SENet. ECANet uses a one-dimensional convolutional layer directly following Global Average Pooling (GAP) and removes the fully connected layer in SENet.
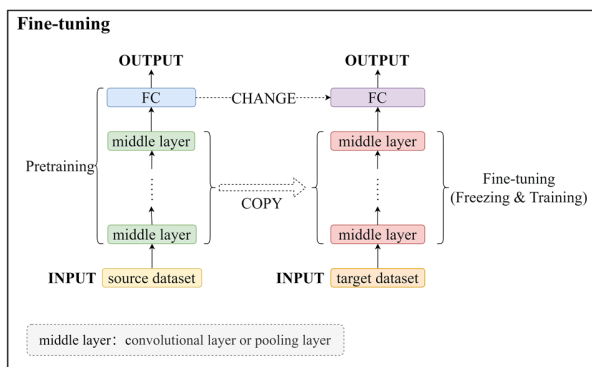

Figure 6: Structure of the fine-tuning approach. The model parameters for extracting generic features trained on the source dataset are frozen, and the unfrozen parameters are re-trained on the target dataset for optimization. The fully connected layer is adjusted based on the number of classification categories.
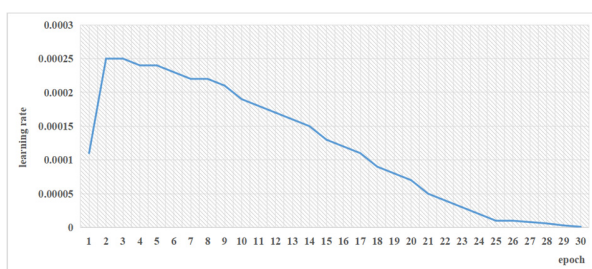

Figure 7: Learning rate variation. The learning rate changes continuously as the epoch increases, the first epoch is the warm-up phase, and the learning rate is tiny. From the second epoch, the learning rate is adjusted to the pre-set value (2.5e-4), and the learning rate is constantly changed based on the Cosine Learning Rate Decay method. The final learning rate was close to 1e-6.

specific structure is shown in Figure 6, and we only needed a small learning rate and a small number of training epochs to achieve a good training performance.

5. Experiments and Analyses

This section describes and analyzes the experimental process of the rice leaf disease identification model in detail and compares the experimental results with other SOTA models.

5.1 Experimental Setup

The experiments of the model were conducted using a high-performance computer (HPC) with the Ubuntu 20.04 LTS operating system. The HPC is equipped with seven Nvidia RTX3090s, each with a memory size of 24 GB, thus having a total video memory of 168 GB. The details of the HPC are shown in Table 2. Pytorch 1.12 was used to implement the proposed method.

To avoid overfitting during model training, the model parameters are divided into two parts: weights and bias. The decay operation was only applied to weights, thus regulating the effect of model complexity on the loss function [18]. In addition, we chose AdamW as the optimizer for ECA-ConvNeXt to get a better training result. To ensure the stability of the model, we used the warm-up method to select a small learning rate at the beginning of training for one epoch. After the learning rate was updated to a pre-set value, the model used the Cosine Learning Rate Decay method to adjust the learning rate. Due to the incorporation of the fine-tuning method, the required learning rate and the number of training epochs are both low. After experimental testing, the training epoch is 30, the initial learning rate is 2.5e-4, and the final learning rate is 1e-6. The learning rate of each epoch is determined jointly by the initial and final learning rate. We confirm that 30 epochs achieved the best accuracy on the testing
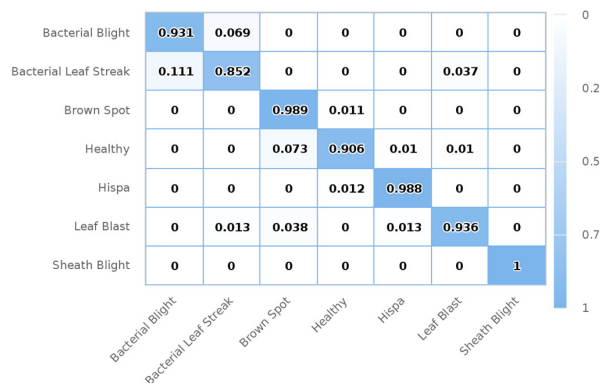
Figure 8: Confusion matrix of the identification results. The confusion matrix is normalized over the true (rows) condition for our proposed model. The horizontal axis represents the predicted classes and the vertical axis represents the true classes. Each cell element represents the proportion of the number of the predicted class to the total number of the true class. The diagonal elements represent correctly classified outcomes. All other elements represent incorrectly classified outcomes.

Item	Specification
Central Processing Unit	AMD EPYC 7543 32-Core Processor × 2
Graphics Processing Unit	Nvidia RTX3090 [PCIe 24GB] × 7
Operating System	Ubuntu 20.04 LTS
Memory	560 GB
Solid State Drive Storage	200 GB
Programming Language	Python 3.9
Version of CUDA	11.4
Deep Learning Framework	PyTorch 1.12

Table 2: Specification of the HPC

Parameter	Value
Optimizer	AdamW
Batch Size	16
Initial Learning Rate	2.5e-4
Final Learning Rate	1e-6
Weight Decay Rate	5e-1

Table 3: Training parameter settings

Type of Disease	Precision	Recall	F1-Score
Bacterial Blight	0.9000	0.9310	0.9153
Bacterial Leaf Streak	0.8846	0.8519	0.8679
Brown Spot	0.9020	0.9892	0.9436
Healthy	0.9775	0.9063	0.9405
Hispa	0.9756	0.9877	0.9816
Leaf Blast	0.9733	0.9359	0.9542
Sheath Blight	1.0000	1.0000	1.0000
Average	0.9447	0.9431	0.9433

Table 4: Precision, Recall, F1-Score, and their average for each type of leaf in the testing dataset.

dataset with an accuracy of 94.82%. The rule to update the learning rate is shown in Figure 7, and other training parameters are set in Table 3.

5.2 Evaluation Metrics

We evaluated the model using the metrics that included Accuracy, Precision, Recall, and F1-Score, commonly used to evaluate models for classification tasks.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

where TP denotes True Positive, TN denotes True Negative, FP denotes False Positive, and FN denotes False Negative. Accuracy is the most commonly used metric in classification performance evaluation. It intuitively reflects the model's classification performance. F1-Score evaluates the model more comprehensively by considering both the Precision and Recall of the model.

5.3 Performance of ECA-ConvNeXt

Our experimental results showed that the ECA-ConvNeXt achieved a better classification performance to identify six types of typical rice leaf diseases and healthy rice leaves than other SOTA models on the testing dataset, which contains 425 images. ECA-ConvNeXt achieved 94.82% for Accuracy, 94.47% for Precision, 94.31% for Recall, and 94.33% for F1-Score. The Precision, Recall, and F1-Score of the seven categories of rice leaf disease identification results in the testing dataset are shown in Table 4. Meanwhile, the confusion matrix of the testing dataset identification results is in Figure 8, which shows that the identification performance of ECA-ConvNeXt was the highest for Sheath Blight with an accuracy of 100% and the lowest for Bacterial Leaf Streak with an accuracy of 85.2%. The accuracy values of Bacterial Blight, Brown Spot, Healthy, Hispa, and Leaf Blast were 93.1%, 98.9%, 90.6%, 98.8% and 93.6%, respectively.

Comparison experiments were conducted to further verify the effectiveness of the ECA-ConvNeXt model compared to other SOTA models in the classification domain. Table 5 compares the performance of four classification SOTA models on the rice leaf disease identification dataset. The results show that the ECA-ConvNeXt model improves the classification Accuracy by 4.2% and 1.6% compared to VGG-19 and VGG-16. The model also improves the classification Accuracy by 1.6% and 2.3% compared to that of ResNet-34 and ResNet-50.

Model	Accuracy
VGG-19	0.906
VGG-16	0.932
ResNet-34	0.932
ResNet-50	0.925
ECA-ConvNeXt	0.948

Table 5: Comparison of classification performance of four SOTA models and ECA-ConvNeXt on the rice leaf disease identification dataset.

Model	Accuracy	Precision	Recall	F1-Score
ConvNeXt +Fine-tuning	0.9412	0.9419	0.9414	0.9409
ConvNeXt +Fine-tuning+ECA	0.9482	0.9447	0.9431	0.9433

Table 6: Experimental results of ablation experiments.

5.4 Ablation Experiments and Analyses

We designed ablation experiments to investigate the effects of the attention mechanism ECA module of the ConvNeXt network on improvements in Accuracy, Precision, Recall, and F1-Score using the same rice leaf disease identification dataset. The baseline model for this experiment is the ConvNeXt network with the fine-tuning method added. The training strategy, data pre-processing, and data augmentation methods were used in the ablation experiments. The experimental results of the ablation experiments are shown in Table 6.

Adding the attention mechanism ECA module between the convolutional layers and the global average pooling layer of the network improves the feature extraction capability of the model with fewer model parameters, resulting in better classification performance. Accuracy was improved by 0.7% compared to the baseline model, while Precision, Recall, and F1-Score were also slightly improved by 0.28%, 0.17%, and 0.24%, respectively.

6. Discussion

Based on the characteristics of the rice leaf disease identification task, we used the ECA attention module and transfer learning to optimize the network structure. These approaches enabled the model to achieve a higher identification accuracy and a better performance in detecting six categories of typical rice leaf diseases and healthy rice leaves. However, in the production environment, the disease category of an individual rice leave may not be unique, which can vastly increase the difficulty of the model to perform the identification task and lead to the degradation of the model performance. In addition, it is not easy to obtain images with a constant size in the natural environment of farmland as agricultural

workers may use different imaging devices or mobile phones to capture rice leaf images. To deal with this problem, we proposed a pre-processing technique to resize images to a resolution of 224×224. This enables the model to take rice leaf images of different sizes as input.

Furthermore, since there are quantitative variations in different types of images and Bacterial Leaf Streak and Sheath Blight are significantly fewer than other types of rice leaf images, we used different levels of data augmentation to process different categories of images to make the dataset balanced in terms of quantity distribution. For future research, we intend to develop a mobile application for labeling rice leaf images on the spot.

In addition, we intend to expand the rice leaf disease identification dataset by manually collecting images of diseased rice leaves in the farmland. We released 3353 sample images of the rice leaf disease identification dataset: https://www.kaggle.com/datasets/wangxiaoqii/rice-leaf-disease-identification-dataset. The full dataset is available upon request.

In terms of its utility, the ECA-ConvNeXt architecture can fit into modern agricultural production environments driven by IoT (Internet of Things) and big data. With the images captured by cameras in the field, we can quickly identify rice leaf diseases and establish a monitoring system for disease control in agricultural production. Nevertheless, the widespread use of the rice leaf disease identification model based on ECA-ConvNeXt and other deep learning models in the future would require infrastructure and hardware, which would incur economic costs.

7. Conclusion

This paper proposes the ECA-ConvNeXt model based on the latest CNN network ConvNeXt to identify six categories of typical rice leaf diseases and healthy rice leaves. We constructed the rice leaf disease identification dataset that contains real-world images. The ECA-ConvNeXt model improves the identification accuracy and the training performance by fine-tuning, introducing the ECA module, and other effective methods. The accuracy of the model is better than other SOTA methods. Applying the ECA-ConvNeXt model in rice cultivation can potentially decrease yield reductions and economic losses. Based on the similarities of diagnostic methods for diseased leaves of different crops, the proposed model has broad application prospects and can be extended to disease identification of other crops.

References

[1] S. Muthayya, J. D. Sugimoto, S. Montgomery, and G. F. Maberly, "An overview of global rice production, supply, trade, and consumption: Global rice production,

consumption, and trade," *Ann. N.Y. Acad. Sci.*, vol. 1324, no. 1, pp. 7–14, Sep. 2014.

[2] Q. Zhang, "Strategies for developing Green Super Rice," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 42, pp. 16402–16409, Oct. 2007.

[3] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, "Detection of rice plant diseases based on deep transfer learning," *J Sci Food Agric*, vol. 100, no. 7, pp. 3246–3256, May 2020.

[4] Y. Lu, J. Du, P. Liu, Y. Zhang, and Z. Hao, "Image Classification and Recognition of Rice Diseases: A Hybrid DBN and Particle Swarm Optimization Algorithm," *Front. Bioeng. Biotechnol.*, vol. 10, p. 855667, Apr. 2022.

[5] K. Gupta, A. Garg, V. Kukreja, and D. Gupta, "Rice Diseases Multi-Classification: An Image Resizing Deep Learning Approach," in *2021 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain, Dec. 2021, pp. 170–175.

[6] R. Deng *et al.*, "Automatic Diagnosis of Rice Diseases Using Deep Learning," *Front. Plant Sci.*, vol. 12, p. 701038, Aug. 2021.

[7] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," *Environ Monit Assess*, vol. 195, no. 1, p. 120, Jan. 2023.

[8] R. P. Narmadha, N. Sengottaiyan, and R. J. Kavitha, "Deep Transfer Learning Based Rice Plant Disease Detection Model," *Intelligent Automation & Soft Computing*, vol. 31, no. 2, pp. 1257–1271, 2022.

[9] Blake CL and Merz CJ, UCI repository of machine learning database.URL: https://archive.ics.uci.edu/ml/index.php. last accessed Feb, 2023.

[10] G. Latif, S. E. Abdelhamid, R. E. Mallouhy, J. Alghazo, and Z. A. Kazimi, "Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model," *Plants*, vol. 11, no. 17, p. 2230, Aug. 2022.

[11] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, Jun. 2022, pp. 11966–11976.

[12] Rice Leaf Disease Image Samples. URL:https://data.mendeley.com/datasets/fwcj7stb8r/1. last accessed Feb, 2023.

[13] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, Oct. 2021, pp. 9992–10002.

[14] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 5987–5995.

[15] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, Jun. 2018, pp. 7132–7141.

[16] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 11531–11539.

[17] Z. Zhou, J. Shin, L. Zhang, S. Gurudu, M. Gotway, and J. Liang, "Fine-Tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 4761–4772.

[18] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of Tricks for Image Classification with Convolutional Neural Networks," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 558–567.