RESEARCH ARTICLE

A High-Quality Rice Leaf Disease Image Data Augmentation Method Based on a Dual GAN

ZHAO ZHANG [1], QUAN GAO[1,2], LIRONG LIU[1], AND YUN HE [1,2]

[1]School of Big Data, Yunnan Agricultural University, Kunming, Yunnan 650201, China
[2]Key Laboratory for Crop Production and Intelligent Agriculture of Yunnan Province, Kunming, Yunnan 650201, China

Corresponding author: Yun He (heyun@ynau.edu.cn)

ABSTRACT Deep learning models need sufficient training samples to support them in the training process; otherwise, overfitting occurs, resulting in model failure. However, in the field of smart agriculture, there are common problems, such as difficulty in obtaining high-quality disease samples and high cost. To solve this problem, this paper proposed a high-quality image augmentation (HQIA) method for generating high-quality rice leaf disease images based on a dual generative adversarial network (GAN). First, the original samples were used to train Improved Training of Wasserstein GANs (WGAN-GP) to generate pseudo-data samples. The pseudo-data samples were put into the Optimized-Real-ESRGAN (Opt-Real-ESRGAN) to generate high-quality pseudo-data samples. Finally, the high-quality pseudo-data samples were put into the disease classification convolutional neural network, and the effectiveness of the method was verified by indicators. Experimental results showed that this method can generate high-quality rice leaf disease images, and the recognition accuracy of high-quality rice disease image samples augmented by this method was 4.57% higher than that of using only the original training set on ResNet18 and 4.1% higher on VGG11. Compared with the data augmentation method only by WGAN-GP, the accuracy of ResNet18 increased by 3.08%, and the accuracy of VGG11 increased by 3.55%. The results demonstrate the effectiveness of the proposed method with limited training datasets.

INDEX TERMS Rice leaf disease, data augmentation, generative adversarial networks, deep learning, image super-resolution.

I. INTRODUCTION

Rice is the main food crop for nearly 50% of the world's population [1]. Disease is one of the main factors that directly causes a large reduction in rice production, which can lead to a reduction of 40–50% or even no grain yield in severe cases [2]. Therefore, timely detection and determination of disease types is of great significance to ensure rice production. The diagnosis and prevention of frequent rice

The associate editor coordinating the review of this manuscript and approving it for publication was Sudhakar Radhakrishnan [ID].

diseases are usually based on characteristic disease symptoms, requiring field knowledge and experience. For non-professional researchers, especially farmers, who are not familiar with the occurrence period and symptoms of rice diseases, artificial misjudgment often occurs, efficiency is low, and expert dependence is high, making it difficult to accurately and timely treat the disease, resulting in rice production reduction [3].

To solve this problem, domestic and foreign scholars began to use automatic image recognition technology for crop disease diagnosis very early. The earliest date can be traced back

to 1973, when Indian researchers used remote sensing technology to obtain rice field images near Mumbai and pattern recognition technology to identify rice bacterial blight [4]. Later studies focused on using pattern recognition technology or ma-chine learning technology to classify processed images of rice plant leaves or panicles for disease image recognition in rice. Among them, the representative studies include: Liu et al. [5] applied a support vector machine (SVM) for the classification and recognition of 15 major diseases on rice leaves; Charlie Paul et al. [6] realized the extraction, classification, and recognition of color, shape, and other features of disease images based on feature matching (FM) technology; Anthony et al. [7] also performed disease image recognition of rice by adopting the membership function (MF) technology. This type of research can be divided into two steps. First, image processing technology is used to extract the feature information of rice images, and then the extracted features are classified and recognized through pattern recognition or machine-learning technology. As image features are the basis of subsequent recognition steps, the quality of features has a crucial impact on the generalization performance of the method. In the early research on disease image recognition, which features to extract and in which way to extract features were determined manually. Therefore, researchers are required to have a high level of domain knowledge. Moreover, when there are many kinds of dis-eases, the relationship between image features is extremely complex, making it difficult to distinguish effective features and set feature weights, which inevitably limits the performance of disease image recognition methods to a certain extent.

In recent years, deep learning technology has developed rapidly, especially in the fields of image processing and natural language processing, and has achieved many results. Unlike traditional machine learning methods, deep learning learns data features through a large sample dataset, can automatically extract and represent the features of input data (images or text), and can describe the intrinsic information of data more accurately and richly. In many fields, deep learning has shown better performance than traditional methods. In the research field of rice disease image recognition, some excellent research results have also emerged in recent years. For example, Liu et al. [4] from the Chinese Academy of Agricultural Sciences realized the recognition of rice leaf blight images based on convolutional neural networks (CNNs), and the accuracy rate was over 90%. Huang et al. [8] used the GoogleNet model of deep CNNs to detect rice blast disease, and the prediction accuracy reached 92%. Tan et al. [9] used a deep CNN to realize the image recognition of eight rice diseases, and the accuracy rate was between 54% and 93%. Shreyasi et al. [10] applied CNNs to achieve 94% accuracy in the recognition of individual rice leaf diseases. The proposed method achieved an accuracy of 78.44% for multi-type rice leaf disease recognition. These studies are based on deep learning models, which realize automatic feature extraction of rice disease images and improve the accuracy of recognition to a certain extent.

Due to the deep network structure and strong expression ability compared with traditional models, the model constructed by the deep learning method requires a large balanced sample dataset as training support; otherwise, it will lead to overfitting [11]. Overfitting means that the model is very effective in training data, but the effect is greatly reduced when encountering test data, invalidating the model at this time. Scholars from Cambridge University and Carnegie Mellon University discussed the number of training samples required for the CNN model commonly used in the depth learning method through hypothesis, proof, and verification [12], and concluded that when the input data is d-dimensional, the m-dimensional convolutional filter is activated linearly to achieve prediction error ε. The number of samples required is $\tilde{O}(m/\varepsilon^2)^1$; the number of samples required for training the full connection layer should be at least $\Omega(d/\varepsilon^2)$. To learn a linear activated single hidden layer CNN model (the output weight dimension is r), the number of training samples required should be $\tilde{O}m + r)/\varepsilon^2)$. However, in practical applications, researchers often face the problem of scarce sample data [13]. The number of samples is too small, or the number of samples of each type is unbalanced, so an effective deep model cannot be trained, but it is difficult and costly to collect enough effective samples. For disease image recognition based on individual leaves, it is necessary to collect enough training samples by shooting leaf by leaf, and the acquisition cost is difficult to predict. Moreover, in the face of a multi-category disease image recognition problem, to ensure the balance of training samples, it is also necessary to ensure that the number of samples of each type of disease is consistent, making the cultivation of diseased plants and the collection of images extremely expensive in time, capital, and labor costs.

In response to these problems, this paper proposes a high-quality image augmentation (HQIA) method for generating high-quality rice leaf disease images based on a dual generative adversarial network (GAN). This method first uses the image generation GAN network to generate the initial image and then uses the image super-resolution GAN network to enhance the image quality of the generated initial image. This method combines image generation and image super-resolution GAN networks and finally generates high-quality rice leaf disease images.

The contributions of this paper include the following three aspects:

1) The proposed method used two GAN networks to generate high-quality images of rice leaf disease. The first GAN network, WGAN-GP, was responsible for generating the initial image, and the second GAN network, an optimized Real-ESRGAN network (named Opt-Real-ESRGAN), was responsible for super-resolution enhancement of the generated image to achieve image clarity improvement. Through the combination of the two GAN networks, high-quality rice disease image samples were generated, and the fine-grained features

of the disease image were generated to the maximum extent.

2) When applying the Opt-Real-ESRGAN network for super-resolution enhancement, the network structure was optimized, network stability was enhanced, the image super-resolution generation speed was accelerated, and the generated samples were more suitable for rice disease image recognition.

3) Experiments were carried out on the real dataset and the image sample dataset generated by the proposed method using the VGG11 and ResNet18 models, which are commonly used in the field of disease image recognition research. The effectiveness of the image data samples generated by the proposed method was verified. These samples effectively expanded the training sample dataset of the disease image model, thereby improving the performance of the disease recognition model.

II. RELATED WORKS

With the wide application of deep learning, scholars are increasingly using the CNN model commonly used in deep learning methods for the identification of crop diseases and pests, but the identification of diseases and pests requires a large number of datasets for model training. However, most existing datasets in the research field of diseases and pests have a small sample size. The unbalanced distribution of disease samples and the collection and labeling process are time-consuming and have a high economic cost, mainly because of the lack of and high cost of dataset samples, which hinders research in this direction. Therefore, data augmentation technology has begun to appear. Data augmentation occurs mainly with the existing small datasets through various algorithms, so that the dataset is expanded into a sufficient number of reasonable and balanced data samples.

Traditional disease image recognition research usually uses simple image transformation algorithms for data augmentation in cases of insufficient samples. For example, Sun et al. [14] used scale transformation and other methods to balance the number of samples of each plant disease. Huang et al. [8] used random color transformation to expand the number of rice disease samples. Ma et al. [15] used rotation transformation and other methods to expand the cucumber disease dataset. Bao et al. [16] expanded the apple disease dataset by adding noise. These data augmentation methods, such as rotation transformation, mirror transformation, scaling transformation, translation transformation, scale transformation, color contrast transformation, and noise transformation, are based on simple changes of the original single image sample, and the difference between the generated image and the original image is small, which cannot improve the diversity and generalization ability of the dataset. In 2014, Goodfellow et al. [17] proposed GANs, which can be divided into two parts: generator and discriminator. The generator obtains the potential distribution of real data samples and generates new data samples. The discriminator is responsible for determining whether the input is real data or generated data, and the two compete with each other to finally reach a Nash equilibrium [18]. To improve the problems existing in traditional image augmentation, researchers began to introduce GAN into the field of crop disease image recognition for data augmentation. Hu et al. [19] compared the accuracy of model recognition of different data augmentation methods on the tea dataset: C-DCGAN [20], traditional data augmentation, and no data augmentation. The results showed that traditional data augmentation improved the accuracy of the model, and the accuracy results of the C-DCGAN model based on GAN were better than those of traditional data augmentation, demonstrating that GAN effectively improved the diversity and generalization ability of the dataset and the accuracy performance. Subsequently, many researchers have conducted similar verifications. For example, Wu et al. [21] used DCGAN [22] to verify the tomato disease dataset in 2020. Bird et al. [23] used conditional GAN [24] to validate a fruit dataset. By learning the overall distribution of the dataset samples and generating an extended dataset consistent with the sample distribution of the original dataset, GAN can avoid the problems of insufficient generalization ability based on a single image change in traditional data augmentation.

Crop disease recognition is essentially fine-grained. Unlike ordinary image classification tasks, the signal-to-noise ratio of fine-grained images is very small, and information containing sufficient discrimination often only exists in a very small local area. High-quality datasets can bring higher recognition rates to disease recognition models [25]. Existing research on GAN-based disease image data augmentation is as follows: Zeng et al. [26] based crop disease image sample augmentation on CycleGAN, and Rongcheng et al. [27] based plant disease image sample augmentation on DCGAN. These disease image sample augmentation studies generally have problems, such as a small image size and poor image quality, which obviously do not meet the requirements of high-quality datasets. To some extent, these issues limit the performance of existing disease image recognition models.

Wang et al. [28] proposed a deep blind image super-resolution model (Training Real-World blind super-resolution with Pure Synthetic Data, Real-ESRGAN). The transformation from a low-quality image to a high-quality image in a real scene is realized. Inspired by this, this paper attempted to use image super-resolution technology to improve the existing GAN image generation algorithm, which has the problems of a small image size and poor image quality. Image super-resolution (SR) is the process of using algorithms to restore degraded images from low resolution (LR) to high resolution (HR), which is an important technology in computer vision and imaging. It has a wide range of applications in medical imaging, image processing, satellite remote sensing, and other fields. For example, in 2018, Gao et al. [29] used DR-GAN to perform super-resolution reconstruction of low-resolution medical images, and the reconstructed high-definition medical influence helped doctors make the correct diagnosis. In 2019, Li et al. [30] used ISRGAN
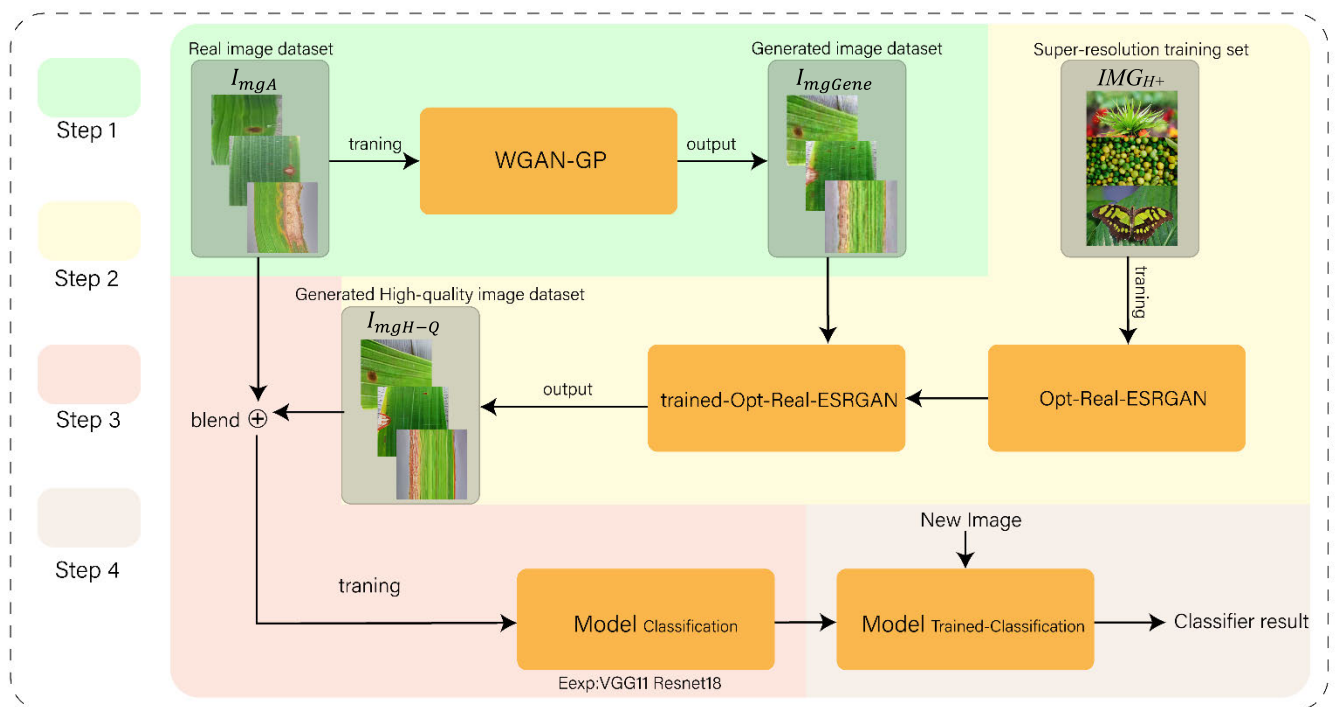
FIGURE 1. Overview of the high-quality image augmentation (HQIA) method.

to perform super-resolution reconstruction of remote sensing images, and the reconstructed high-definition satellite images contained more information, improving meteorological monitoring and geological exploration. However, there is no related research in the field of crop disease image recognition.

To solve the problems of the scarcity of existing crop disease image recognition samples and the low quality of sample images generated by existing image augmentation methods. In this paper, the HQIA method was proposed for generating high-quality rice leaf disease based on a dual GAN. This method uses the data augmentation GAN network to generate a batch of pseudo data samples, and uses the super-resolution GAN network to enhance the generated pseudo data samples. Through this method, higher quality disease image samples were generated compared with traditional research, expanding the training space of the disease recognition model and improving recognition performance.

III. METHODOLOGY
A. OVERVIEW OF THIS METHOD
In this paper, we proposed the HQIA method, a dual-GAN-based method for high-quality rice leaf disease generation. This method used a WGAN-GP network to generate image samples, and then used an optimized Real-ESRGAN (named Opt-Real-ESRGAN) network for super-resolution to enhance the generated image samples. The basic framework is shown in Fig. 1. In this chapter, the implementation process and implementation of HQIA are introduced in detail.

The basic implementation process of HQIA was as follows:
1) The real rice disease image sample dataset I_{mgA} was inputted into WGAN-GP for training, and the pseudo image data sample set I_{mgGene} was generated.
2) IMG_{H+} training Opt-Real-ESRGAN was used to obtain the trained-Opt-Real-ESRGAN. I_{mgGene} was inputted to the trained-Opt-Real-ESRGAN generator, and the output I_{mgH-Q} had super-resolution.
3) The dataset I_{mgH-Q} was inputted into the image recognition model for training, and the trained image recognition model $Model_{Trained-Classification}$ was obtained.
4) When a new disease image appeared, the image was inputted into $Model_{Trained-Classification}$ to obtain image recognition results.

B. DATA SAMPLE GENERATION BASED ON THE WGAN-GP
The first step of HQIA was to input the real disease sample dataset I_{mgA} into the WGAN-GP. Through iterative training of the generator and discriminator, a trained
WGAN-GP generator was used to generate the initial data augmented sample I_{mgGene}.

1) NETWORK ARCHITECTURE
As shown in Fig. 2, the WGAN-GP based on a neural network was divided into two parts: generator (G) and discriminator (D). The general learning process was as follows: as the input of G, z was noise, which could be Gaussian noise, usually uniform noise. G generated new data samples with given noise by learning the mathematical distribution state of real
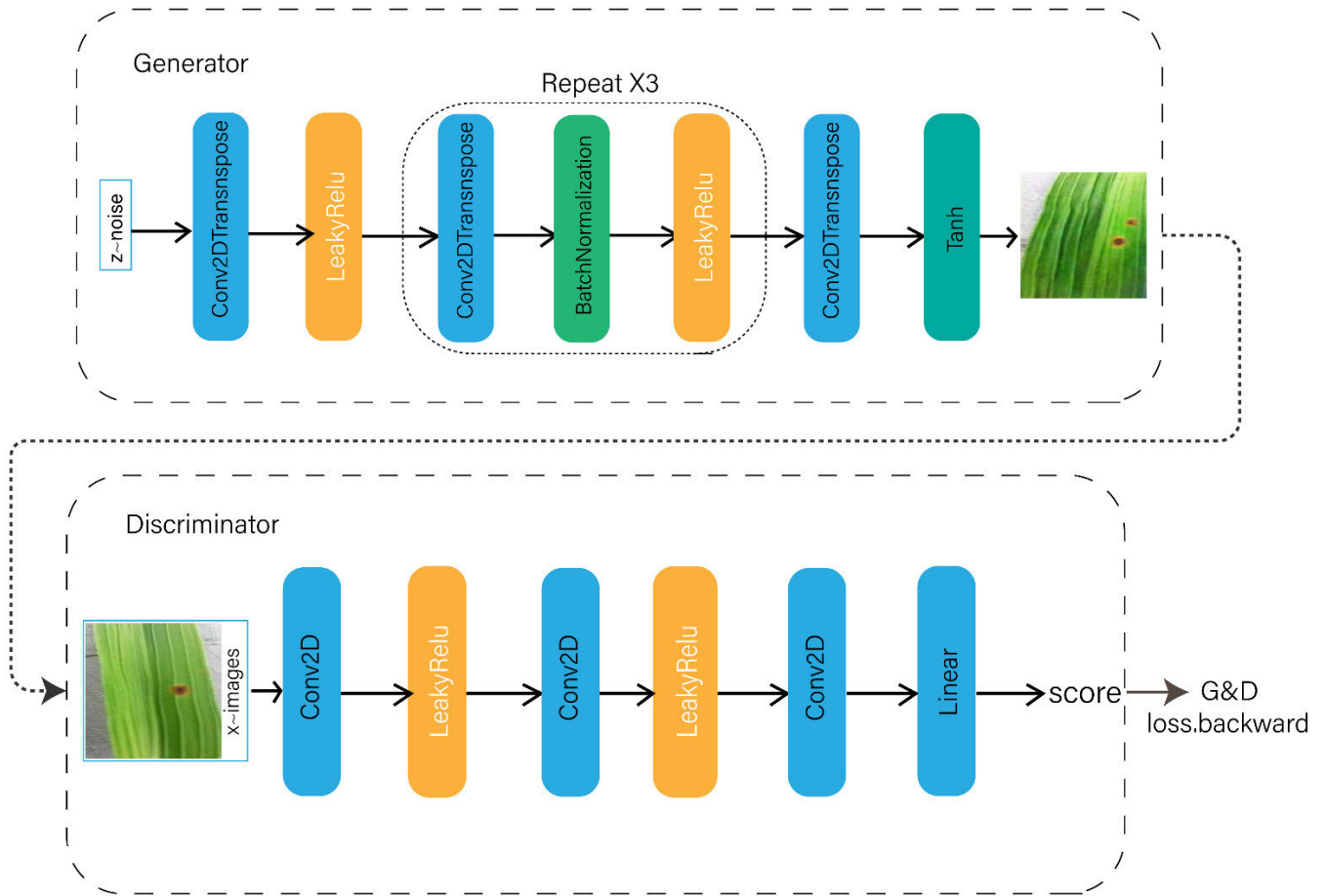
FIGURE 2. WGAN-GP's generator (top) and discriminator architecture (bottom).

data samples. D also learned the mathematical distribution state of real data samples and then identified the data samples generated by G as real or generated data samples. According to the judgment results of D, D, and G improved their model parameters. They competed with each other and eventually reached Nash equilibrium.

2) OBJECTIVE FUNCTION

WGAN-GP was selected as the data augmentation network for HQIA. WGAN-GP was based on WGAN to increase the gradient penalty mechanism. To better understand WGAN-GP, we first introduced the objective function of the WGAN [31]. This is shown in Eq. 1 below:

$$\text{WGAN} = \frac{min}{G} \frac{max}{D \in \mathcal{D}} = \mathop{\mathbb{E}}_{x \sim \mathbb{P}_r}[D(x)] - \mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_g}[D(\tilde{x})], \quad (1)$$

where D is a set of 1-Lipschitz [32] functions, x is real data or real pictures, $\mathbb{E}$ is mathematical expectation, $\mathbb{P}_r$ is the probability distribution of x, $\tilde{x}$ is the random noise of the input, and $\mathbb{P}_g$ is the probability distribution of $\tilde{x}$.

WGAN is prone to the problem of gradient explosion. Thus, the gradient penalty GP term [31] was added to the

discriminator loss function of the WGAN-GP network instead of weight clipping to enforce Lipschitz constraints to avoid gradient explosion. The objective function of the gradient penalty mechanism is given as Eq. 2 below:

$$\text{GP} = \lambda \mathop{\mathbb{E}}_{\hat{x} \sim \mathbb{P}_g} \left[(||\nabla_{\hat{x}} D(\hat{x})||_2 - 1)^2 \right], \quad (2)$$

where $\hat{x}$ represents random interpolation sampling on the line connecting $\mathbb{P}_r$ and $\mathbb{P}_g$.

The WGAN-GP network objective function after introducing the gradient penalty mechanism is shown in Eq. 3:

$$\text{WGAN} - \text{GP} = \underbrace{\mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_g}[D(\tilde{x})] - \mathop{\mathbb{E}}_{x \sim \mathbb{P}_r}[D(x)]}_{\text{Original critic loss}}$$
$$+ \underbrace{\lambda \mathop{\mathbb{E}}_{\hat{x} \sim \mathbb{P}_g} \left[\left(\left\| \nabla_{\hat{x}} D(\hat{x}) \right\|_2 - 1 \right)^2 \right]}_{\text{Our gradient penalty}}. \quad (3)$$

C. IMAGE SAMPLE SUPER-RESOLUTION ENHANCEMENT BASED ON OPT-REAL ESRGAN

The second step of HQIA was to adopt the Opt-Real-ESRGAN network for the super-resolution of the generated
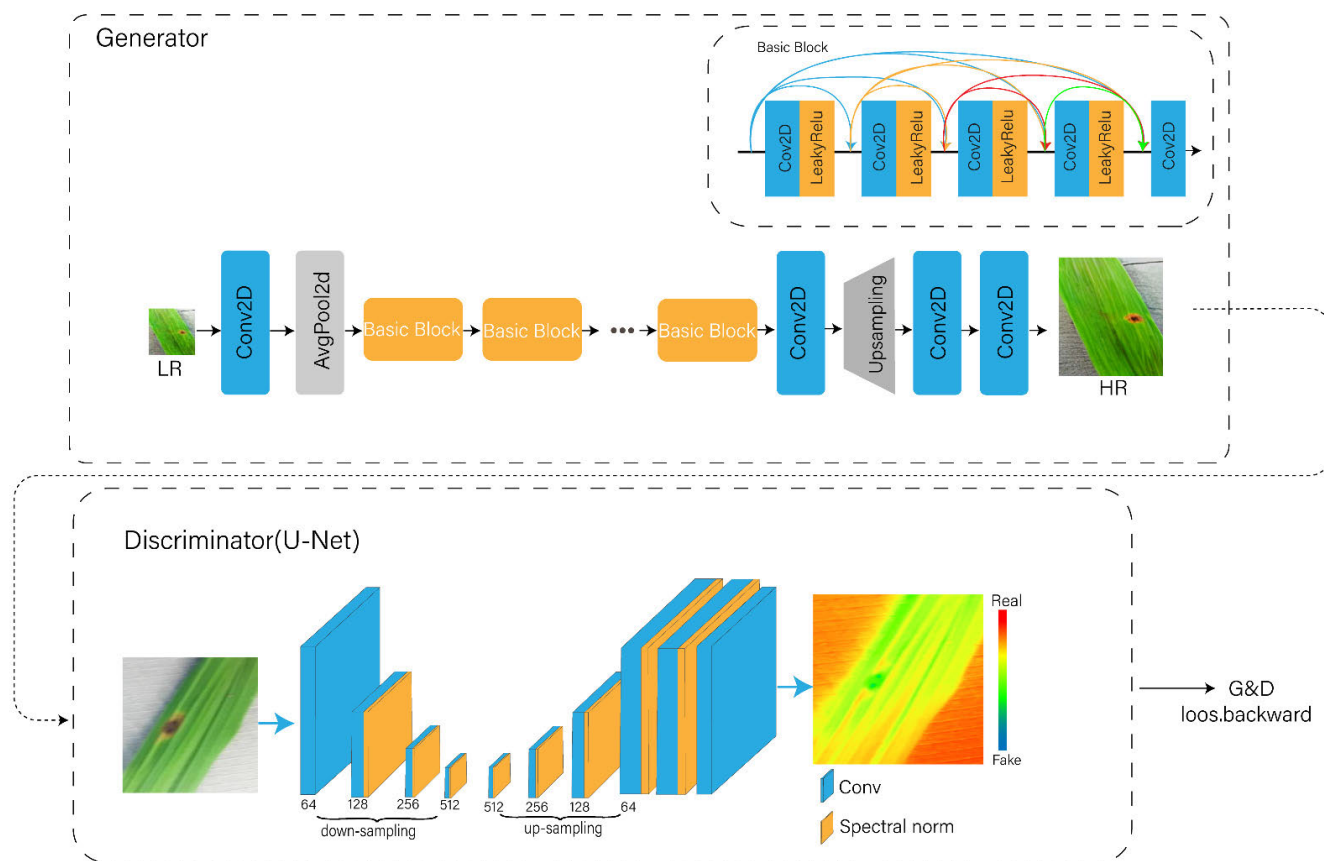
FIGURE 3. Opt-Real-ESRGAN network structure diagram.

image. The super-resolution training dataset IMG_{H+} was used to train Opt-Real-ESRGAN, and the trained-Opt-Real-ESRGAN model was finally obtained through continuous iterative training. The data augmentation sample I_{mgGene} obtained in the previous step was inputted into the generator network of the trained-Opt-Real-ESRGAN model to generate the high-quality disease dataset I_{mgH-Q}.

1) NETWORK ARCHITECTURE

Opt-Real-ESRGAN was divided into two parts: G and D. According to the mapping relationship between the high-resolution image and the low-resolution image, G used the input low-resolution image through a series of operations, such as Conv2D, AvgPool2d, LeakyRelu, and Upsampling, and finally output the high-resolution image. D was the U-Net discriminator based on spectral normalization (SN). It differed from the traditional discriminator in that it needed to produce gradient feedback of accurate local texture rather than simply judging the global style. The output was the probability that each pixel was correct and provided detailed pixel-by-pixel feedback to the generator. Finally, results G and D were fed back and optimized according to the judgment of D.

The Opt-Real-ESRGAN network structure designed in this paper is shown in Fig. 3. Compared with the Real-ESRGAN

model designed by Wang et al. [2], the Opt-Real-ESRGAN model in this paper added a AvgPooling layer before the Basic Block module of the generator because the Real-ESRGAN model was too large. This costs a lot of time and machine performance in image super-resolution. Increasing the Avg-Pooling layer reduced the size of the image, enhanced the generalization ability, reduced the number of neurons, and reduced the calculation amount of the network model while retaining the characteristics of the original image. We verify this statement in the experimental section.

2) IMAGE DEGRADATION

In the process of image generation, storage, processing and transmission, the imperfection of the imaging system, recording device, transmission medium, and processing method lead to the degradation of image quality, called image degradation. The image super-resolution network realized the restoration of images from low to high resolution by learning the mapping relationship between high-resolution images and degraded low-resolution images. The image super-resolution network realized the restoration of images from low to high resolution by learning the mapping relationship between high-resolution images and degraded low-resolution images. The image degradation process of Opt-Real-ESRGAN mainly included four steps:
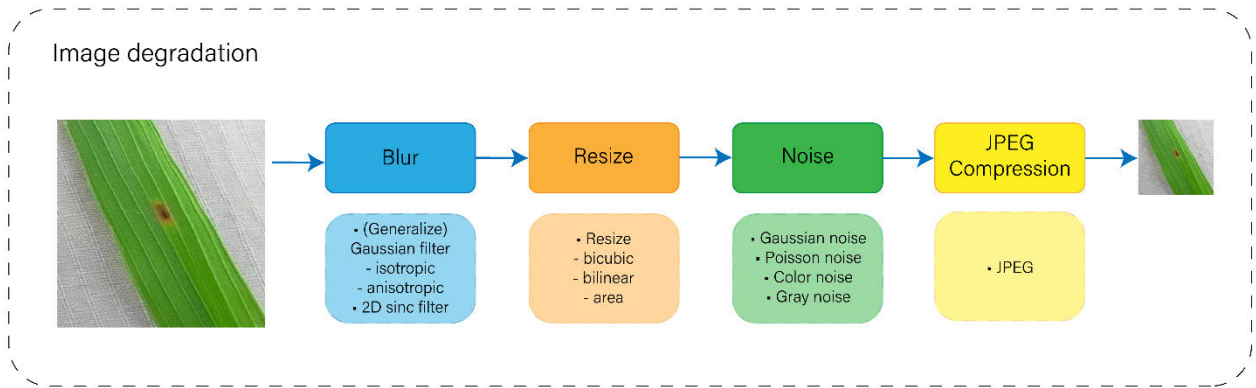
FIGURE 4. Image degradation process.

Blur, Resize, Noise, and JPEG Compression. Blur used Gaussian and 2D sinc filters. The down-sampling factors used by Resize were bicubic, bilinear, and area. Noise was used for Gaussian noise, Poisson noise, Color noise, and Gray noise, and Compression was JPEG Compression. Opt-Real-ESRGAN uses a high-order degradation model. A high-order degradation model means that the image is degraded n times (Eq. 4), whereas the classical degradation model (Eq. 5) is used for each degradation process. However, Blur, Resize, Noise, and JPEG Compression were random at each step.

$$x = D^n(y) = (Dn \circ \cdots \circ D_2 \circ D_1)(y). \tag{4}$$
$$x = \mathcal{D}(y) = \left[(y \circledast k) \downarrow_r + n\right]_{JPEG}, \tag{5}$$

where y is the original image, k is the blur kernel, $\downarrow_r$ is the down-sampling factor, n is the noise, and JPEG is the JPEG compression. These degradation modes are widely used in real-world images. One degradation process of the image is shown in Fig. 4.

3) OBJECTIVE FUNCTION
The graph super-resolution GAN network is different from the image generation GAN network, which needs to perceive the subtle differences between low-resolution images and high-resolution images. In the Real-ESRGANS network, L1 loss, perceptual loss, and GAN loss were used. The objective function is defined as in Eq. 6:

$$L_{total} = w_1 L_{perceptual} + w_2 L_1 + w_3 L_{GAN}, \tag{6}$$

where $L_{perceptual}$ represents Perceptual loss, L_1 represents L1 loss, L_{GAN} represents GAN loss, and w_1, w_2, w_3 represent the weights corresponding to their respective loss functions.

$L_{perceptual}$ uses this loss function to improve the perceptual quality of the image and reconstruct the texture and details of the image. The objective function is defined as in Eq. 7:

$$L_{perceptual} = \frac{1}{C_i H_i W_i} \|D_i(G(x)) - D_i(y)\|_2^2, \tag{7}$$

where D_i is the ith layer of the discriminator subnetwork, C_i is the number of channels corresponding to the ith layer,

H_i and W_i are the length and width of the ith layer feature map, respectively.

The L_1 loss function is also a commonly used loss function in super-resolution networks, which can calculate the L_1 distance between the output image and the real image. The objective function is defined as in Eq. 8:

$$L_1 = E_{I_{LR}} \| G(I_{LR}) - I_{HR} \|_1, \tag{8}$$

where I_{LR} is the low-resolution image, and I_{HR} is the high-resolution image.

L_{GAN} recovered the high-frequency detail information in the low-resolution image so that the generated high-resolution image passed the discrimination of the discriminator with the maximum probability. The objective function is defined as in Eq. 9:

$$L_{GAN} = \begin{matrix} min & max \\ G & D \end{matrix} = [log\ (D)] + [log\ (1 - D(\tilde{x}))], \tag{9}$$

where G is the generator, D is the discriminator, and $\tilde{x}$ is the super-resolved image.

D. TRAINING IMAGE RECOGNITION MODELS ON REAL-GENERATED DATA
The I_{mgH-Q} disease data sample generated in the previous step was merged with the real disease data sample I_{mgA}, and the merged disease dataset was sent to the disease classification model $Model_{Classification}$. After training, we finally obtained the trained disease classification model $Model_{Trained-Classification}$. The process is shown in Fig. 5.
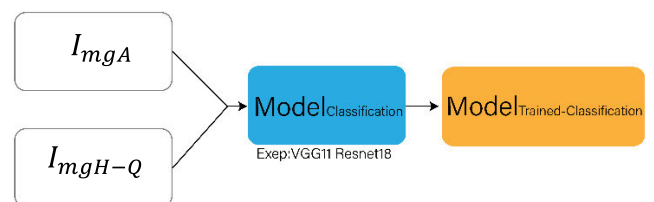

FIGURE 5. Disease classification model training.

E. DISEASE IMAGE RECOGNITION

When a new disease image was obtained, the disease image was inputted into $\text{Model}_{\text{Trained-Classification}}$ in the previous step, and the model output was the disease type of the image. The process is shown in Fig. 6.
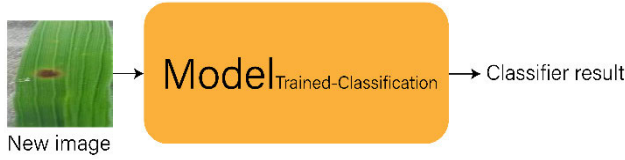

FIGURE 6. Recognition of disease images.

IV. EXPERIMENTS

A. EXPERIMENTAL DATASET

Three representative rice diseases were selected for the dataset in this paper: brown spot, leaf blast, and bacterial blight. Brown spot and leaf blast disease datasets were previously used in [33], and bacterial blight disease datasets were obtained from the Hefei Institute of Intelligent Machinery, Chinese Academy of Sciences [34]. Fig. 7 is an example diagram of these three diseases. To avoid the poor diversity of model recognition caused by the different numbers of samples of each disease, the number of samples of each disease type was balanced, and 846 rice disease leaf samples were selected for each disease, and a total of 2538 rice disease leaf samples were selected. The images were stored in jpg format, and the size was normalized to 300∗300 pixels. Table 1 shows the basic situation of the dataset.

TABLE 1. The rice leaf dataset.

Disease categories	Number of images	Training set: Testing set:
Brown Spot	846	
Leaf Blast	846	50%:50%
Bacterial Blight	846	

FIGURE 7. Leaf disease image samples in the dataset.

The disease dataset I_{mg} was divided into a training set I_{mgA} and test set I_{mgB} with a 1:1 ratio, and the relationship between

the two was as follows:

$$I_{mg} = (I_{mgA}, I_{mgB}). \tag{10}$$

$$I_{mgA} \cap I_{mgB} = \emptyset. \tag{11}$$

$$I_{mgA} \cup I_{mgB} = I_{mg}. \tag{12}$$

B. PERFORMANCE METRICS

Many researchers use accuracy, precision, recall, and F1-measure as indicators to evaluate classification performance and select the best round for evaluation. In this paper, they were selected as criteria for the recognition performance of disease images. They are defined as follows:

$$\text{Accuracy} = \frac{tp + tn}{tp + fp + tn + fn}, \tag{13}$$

$$\text{Precision} = \frac{tp}{tp + fp}, \tag{14}$$

$$\text{Recall} = \frac{tp}{tp + fn}, \tag{15}$$

$$\text{F1} - \text{measure} = \frac{2 \times Precision \times Recall}{Recall + Recall}, \tag{16}$$

Accuracy (Eq. 13) is the ratio of the number of correctly classified predictions to the total number of predictions; generally speaking, the higher the accuracy, the better the classifier. However, in the case of binary classification and unbalanced positive and negative samples, especially when the class is interested in a small number of samples, the referability of the accuracy index is reduced. Precision (Eq. 14) is relative to the prediction of the model. This can be understood as the confidence score of the model when making a new prediction or the likelihood that the prediction is correct. Recall (Eq. 15) is relative to the real label, which can be understood as the proportion of positive cases predicted by the model in actual positive cases. The F1-measure (Eq. 16) is a metric (harmonic mean) that considers both precision and recall and is mainly used to compare the performance of two classifiers; a higher value is better. In the formula, tp is used to successfully predict a positive sample as positive; tn is used to successfully predict a negative sample as negative; fp is used to incorrectly predict a negative sample as positive; and fn is used to incorrectly predict a positive sample as negative.

In addition, to verify the quality of the super-resolution image, this paper used the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and Picture output speed (POS) as evaluation criteria to measure the super-resolution image.

PSNR (Eq. 17) is a common and widely used objective standard for image super-resolution tasks, but the PSNR score cannot be completely consistent with the visual quality seen by the human eye because the human eye is not absolutely sensitive to errors. The perception result is affected by many factors and changes.

$$\text{PSNR} = 10 \times log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right), \tag{17}$$

In the PSNR formula, MSE is the mean square error between the original image and the processed image.

SSIM (Eq. 18) is an index to measure the similarity of two images. From the perspective of image composition, SSIM defines structural information as independent of brightness and contrast, reflecting the attribute of object structure in the scene, and models distortion as a combination of three different factors: brightness, contrast, and structure. The mean is used as an estimate of brightness, the standard deviation is used as an estimate of contrast, and the covariance is used as a measure of the degree of structural similarity.

$$\text{SSIM}(x, y) = \frac{(2\mu_x \times \mu_y)(2\sigma_{xy} + c_2)}{(u_{x+}^2\mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (18)$$

In the SSIM formula μ_x is the average of x, μ_y is the average of y, σ_x^2 is the variance of x, σ_y^2 is the variance of y, σ_{xy} is the covariance of x and y, $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ are constants used to maintain stability, and L is the dynamic range of pixel values ($k_1 = 0.001$, $k_2 = 0.03$).

POS (Eq. 19) is the time it takes to obtain the super-resolution of a single image. Image super-resolution is a very time-consuming and machine performance-consuming task, and large-scale datasets for super-resolution are sensitive to time performance. To more intuitively react to the performance of the model, we introduced POS.

$$\text{POS} = \frac{N}{T}, \quad (19)$$

In the POS formula, N is the total number of processed pictures, T is the total time of processing pictures, and the time unit is s. The lower the POS, the better the performance.

C. EXPERIMENTAL PROCESS AND ANALYSIS

The method in this paper was verified on the above rice dataset, which contains 3 rice diseases, 2538 images in the I_{mg} dataset, 1264 images in the I_{mgA} dataset, and 1264 images in the I_{mgB} dataset. Four main research questions were clarified and verified in this experiment:

RQ 1: Can the data samples generated based on the WGAN-GP be used as supplementary training for the rice disease recognition model?

RQ 2: Can enhanced data samples based on Opt-Real-ESRGAN further effectively improve the accuracy of image recognition?

RQ 3: Opt-Real-ESRGAN and Real-ESRGAN were compared to determine whether Opt-Real-ESRGAN was more suitable for rice data augmentation tasks.

RQ 4: If only image super resolution technology is used without data augmentation, can the performance achieve the same effect as RQ2?

To answer RQ 1, we used WGAN-GP to generate the dataset I_{mgGene}. Separately, the $I_{mgGene} + I_{mgA}$ dataset was used to train the rice disease recognition model, and the I_{mgA} dataset was used to train the rice disease recognition model. The two disease recognition models were compared and evaluated using the disease image recognition performance standard.

To answer RQ 2, the I_{mgGene} dataset generated in the previous step was inputted into the Opt-Real-ESRGAN model to output the super-resolution dataset I_{mgH-Q}. $I_{mgH-Q} + I_{mgA}$ dataset, which was used to train the rice disease recognition model. The answer to question 1 was compared horizontally to verify the disease recognition performance.

To answer RQ 3, Opt-Real-ESRGAN and Real-ESRGAN were used to enhance the dataset I_{mgGene}, and the evaluation criteria were used to compare the results.

To answer RQ 4, We output the super-resolution data set I_{mgAH-Q} on the I_{mgA} dataset, and used the I_{mgAH-Q} dataset to train the rice disease recognition model, and then conducted a lateral comparison with RQ2 to verify the disease recognition performance.

In the experiment, VGG11 and Resnet18, which are frequently-used in the field of crop disease image recognition, were selected as classification models. Both VGG11 and Resnet18 are typical classification models in recent years. They have different structures, and are widely used in many image recognition research fields. In this paper, two different models are selected for experiment, mainly to verify the universality of the proposed method. In order to ensure the uniformity of the experiment, the disease recognition model (VGG11, ResNet18) in this paper uses the Adam optimizer, the learning rate is 0.001, the batch size is set 64, and the epoch is set 100.

1) IMAGE GENERATION BASED ON THE WGAN-GP

The WGAN-GP model was trained using the I_{mgA} dataset, and the best hyperparameters were determined through multiple experiments, as shown in the following table.

TABLE 2. WGAN-GP hyperparameter.

Hyper-parameter	Parameter Settings
Optimizer	Adam
Learning rate	0.0002
Batch size	64
Epochs	7000
Gradient punishment	10
Generate image dimensions	128*128(Pixel)

WGAN-GP was trained with the above hyperparameter settings, and the rice disease images generated by the WGAN-GP after iterative training are shown in Fig. 8. The image quality reached the best effect after 7000 epochs of iteration, and the improvement of continued training was limited, so the 7000 epochs model was uniformly selected.

Using the generator trained with 7000 rounds of the WGAN-GP model, 1264 images of the I_{mgGene} dataset were generated. The VGG11 and ResNet18 models were trained using the $I_{mgGene} + I_{mgA}$ dataset and the I_{mgA} dataset, respectively. After training, the model was tested using the I_{mgB} dataset, and evaluation indicators were used to evaluate the model. The results are shown in Table 3.
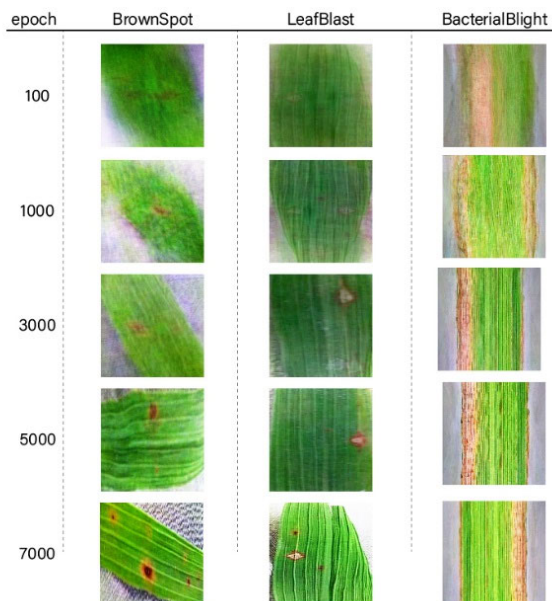
FIGURE 8. WGAN-GP-generated images.

TABLE 3. Rice disease recognition results of the best ResNet18 model.

Performance Metric	I_{mgA}	$I_{mgGene}+I_{mgA}$
Accuracy	87.08%	88.57%
Precision	87.87%	88.73%
Recall	87.10%	88.57%
F1-measure	87.48%	88.65%

Table 3 shows the rice disease identification results of the best ResNet18 model. When using ResNet18 as the disease image recognition model, the accuracy of the trained ResNet18 model using only the I_{mgA} dataset was 87.08%, and the accuracy of the dataset using $I_{mgGene} + I_{mgA}$ improved to 88.57%, an increase of 1.49%. The precision, recall, and F1-measure of the trained ResNet18 model using only the I_{mgA} dataset were 87.87, 87.10, and 87.48%, respectively. When using the $I_{mgGene} + I_{mgA}$ dataset, the recall, precision, and F1-measure were improved to 88.73, 88.57, and 88.65%, respectively, which increased by 0.86, 1.47, and 1.17%, respectively.

Table 4 shows the rice disease identification results of the best VGG11 model. When VGG11 was used as the disease image recognition model, the accuracy of the ResNet18

TABLE 4. Rice disease recognition results of the best VGG11 model.

Performance Metric	I_{mgA}	$I_{mgGene}+I_{mgA}$
Accuracy	86.68%	87.23%
Precision	87.00%	87.80%
Recall	86.70%	87.23%
F1-measure	86.84%	87.52%

model using only the I_{mgA} dataset was 86.68%, and the accuracy of the dataset using $I_{mgGene} + I_{mgA}$ improved to 87.23%, an increase of 0.55%. The precision, recall, and F1-measure of the ResNet18 model using only the I_{mgA} dataset were 87.00, 86.70, and 86.84%, respectively. When using the dataset of $I_{mgGene} + I_{mgA}$, the recall, precision, and F1-measure were improved to 87.80, 87.23, and 87.52%, respectively, which increased by 0.8, 0.53, and 0.68%, respectively.

These experimental results can answer RQ 1. The data samples generated based on the WGAN-GP could be used to supplement the training samples of the disease image recognition model to improve the overall recognition accuracy of the model.

The original disease dataset I_{mg} was divided into a training set I_{mgA} and test set I_{mgB} with a 1:1 ratio. We trained classification models using I_{mgA} and $I_{mgGene} + I_{mgA}$, respectively. Then we use I_{mgB} to test the performance of the classification models. Fig.9 shows the confusion matrix result on original test dataset I_{mgB}. In the ResNet18 confusion matrix of the I_{mgA} training set, the model had poor diversity and predicted many leaf blast samples as bacterial blight samples. This is because the two disease spots have similar colors, and the bacterial blight lesion is larger. Therefore, the disease samples belonging to the leaf blast were predicted to be bacterial blight samples. The ResNet18 confusion matrix of the $I_{mgGene} + I_{mgA}$ training set performed slightly better in terms of model diversity. By increasing the number of training set samples, the model learned the difference between leaf blast and bacterial blight diseases. The VGG confusion matrix of the I_{mgA} training set can confirm the above inference. The diversity of the confusion matrix of the model was also poor, and most leaf blast samples were predicted to be bacterial blight samples. $I_{mgGene} + I_{mgA}$ training set learned the difference between the two diseases by increasing the number of training set samples and increasing the diversity of the confusion matrix. A new problem was also observed in the confusion matrix shown in Fig. 9. Although the leaf blast samples were prevented from being incorrectly predicted as bacterial blight samples by adding disease samples to the training set, the prediction accuracy of the brown spot samples decreased. More brown spot samples were incorrectly predicted as leaf blast samples because brown spot and other disease spots are smaller, and it was easier for the model to learn large and obvious disease spots. However, the image generated by WGAN-GP had more noise, and the imaging quality was relatively poor, as shown in Fig. 10.

2) OPT-REAL-ESRGAN PERFORMS SUPER-RESOLUTION ON THE GENERATED DATA SAMPLES

When training the Opt-Real-ESRGAN model, it is necessary to use high-quality image datasets for training. Similar to Real-ESRGAN [28], this paper used common super-resolution model training datasets DIV2K [35], OST [36], and OutdoorScene [37]. The training details are given in Table 5.
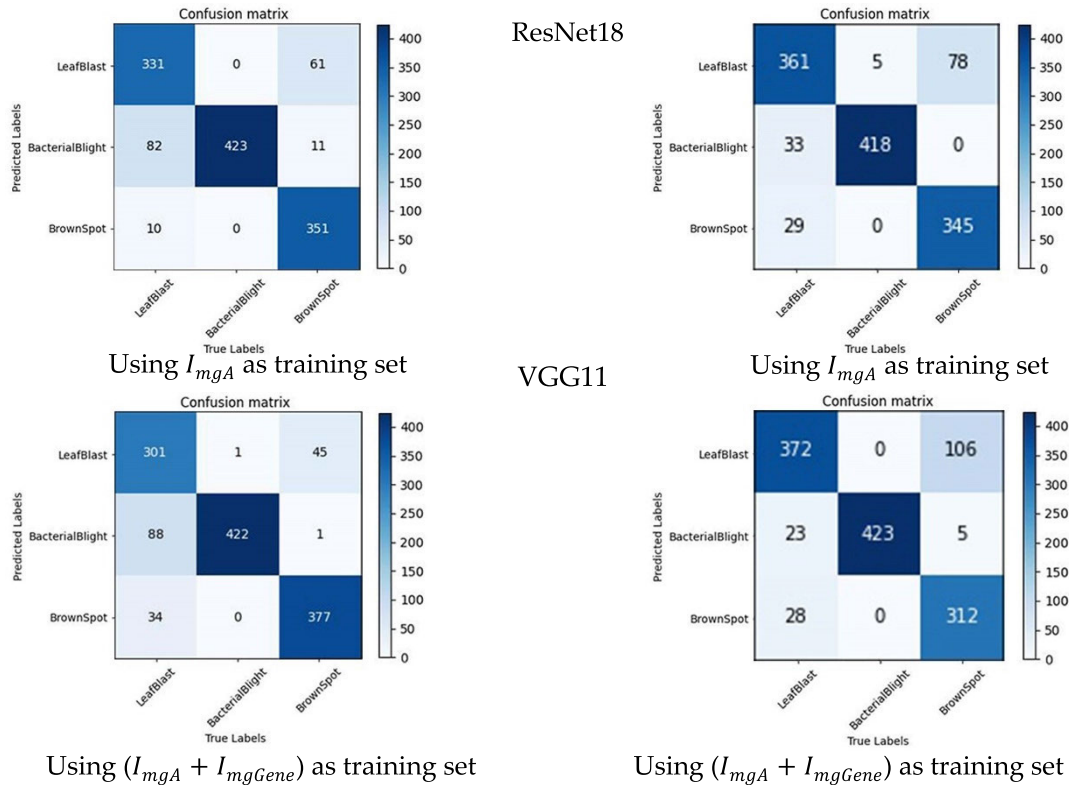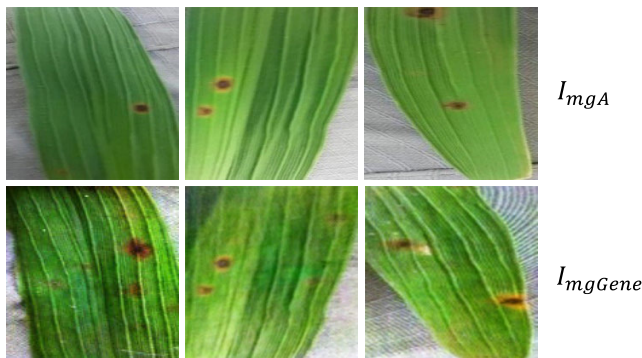
ResNet18

Using I_{mgA} as training set

Using I_{mgA} as training set

VGG11

Using $(I_{mgA} + I_{mgGene})$ as training set

Using $(I_{mgA} + I_{mgGene})$ as training set

FIGURE 9. Confusion matrix on test set.

FIGURE 10. Comparison of samples from **I**mgA and **I**mgGene datasets.

TABLE 5. Real-ESRGAN hyperparameter.

Hyper-parameter	Parameter Settings
Training dataset	DIV2K, OST, OutdoorScene
Loss function Weights	$w_1 = 1;\ w_2 = 1; w_3 = 0.1;$
Optimizer	Adam
Learning rate	1e-4
Weight decay	0
Betas	[0.9, 0.99]
Batch size	48
Epochs	1000K

Opt-Real-ESRGAN was trained using the hyperparameter settings described above. The I_{mgGene} dataset generated in the previous section is input into the trained Opt-Real-ESRGAN generator to generate the x2 upsampled I_{mgH-Q} super-resolution dataset.

Fig. 11 Shows the comparison of imaging effects of Opt-Real-ESRGAN for different diseases. I_{mgA} is the original disease training set samples, and I_{mgGene} is the disease training set samples generated by WGAN-GP. I_{mgH-Q} is the disease training set sample after Opt-Real-ESRGAN super-resolution. The red frame is an enlarged image of the disease spot. In the brown spot disease sample, the background of the I_{mgGene} disease sample was messy, the disease spot part was

fuzzy, and the edge part of the leaf was not sharp enough, while the background of I_{mgH-Q} was clean, the disease spot part was clear, and the edge part of the leaf was sharp. The image quality presented by I_{mgH-Q} was comparable to that of the I_{mgA} original disease training samples. The same problems were observed in leaf blast disease samples. I_{mgGene} disease samples had messy backgrounds, fuzzy disease spots, and leaf edges that were not sharp enough, while I_{mgH-Q} disease samples did not have these problems. Through the enlargement of the disease spot part, the detail was obviously better than the I_{mgGene} disease sample. The bacterial blight I_{mgA} original disease sample had too many disease spot details, while the WGAN-GP model of the I_{mgGene} disease
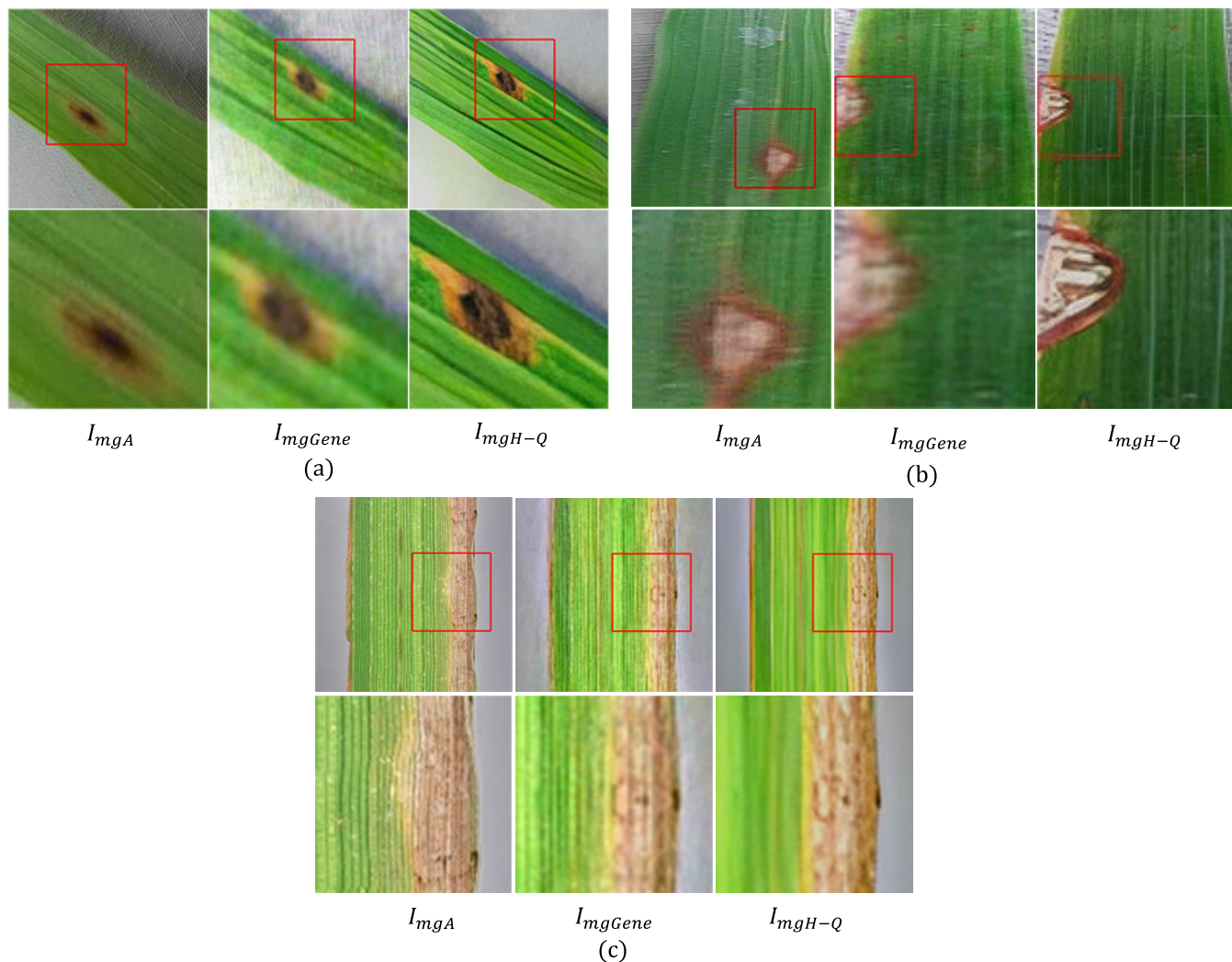
FIGURE 11. Comparison of imaging effects of Opt-Real-ESRGAN with different diseases. (a) Comparison of imaging effects of Opt-Real-ESRGAN for brown spot disease. (b) Comparison of imaging effects of Opt-Real-ESRGAN for leaf blast disease. (c) Comparison of imaging effects of Opt-Real-ESRGAN for bacterial blight disease.

sample had limited learning ability and was not fully learned, limiting the super-resolution ability of Opt-Real-ESRGAN. As a result, the image quality of I_{mgH-Q} and I_{mgA} original disease samples differed, but the I_{mgH-Q} disease training samples formed after the super-resolution of the Opt-Real-ESRGAN model were compared with I_{mgGene}; the background was cleaner, and the edge of the leaf was sharper. The quality of I_{mgH-Q} disease samples was still better than that of I_{mgGene} disease samples.

According to the experimental design, the VGG11 and ResNet18 models were trained by the $I_{mgH-Q} + I_{mgA}$ dataset. After training, the I_{mgB} dataset is used to test the model and the evaluation metrics are used to compare the models.

Table 6 shows the rice disease identification results of the best ResNet18 model. When using ResNet18 as the disease image recognition model, the accuracy using the I_{mgA} dataset was 87.08%, and the accuracy of the ResNet18 model using

TABLE 6. Rice disease recognition results of the best ResNet18 model.

Performance Metric	I_{mgA}	$I_{mgGene}+I_{mgA}$	$I_{mgH-Q}+I_{mgA}$
Accuracy	87.08%	88.57%	91.65%
Precision	87.87%	88.73%	92.00%
Recall	87.10%	88.57%	91.67%
F1- measure	87.48%	88.65%	91.83%

the $I_{mgGene} + I_{mgA}$ dataset was 88.57%. The accuracy of the dataset using $I_{mgH-Q} + I_{mgA}$ improved to 91.65%. Compared with the I_{mgA} dataset, it was 4.57% higher than that using the $I_{mgGene} + I_{mgA}$ dataset, and it was 3.08% higher than that using the $I_{mgGene} + I_{mgA}$ dataset. The precision, recall, and F1-measure of the ResNet18 model using the I_{mgA} dataset were 87.87, 87.10, and 87.48%, respectively. The precision, recall, and F1-measure of the ResNet18 model using the $I_{mgGene} + I_{mgA}$ dataset were 88.73, 88.57, and 88.65%,
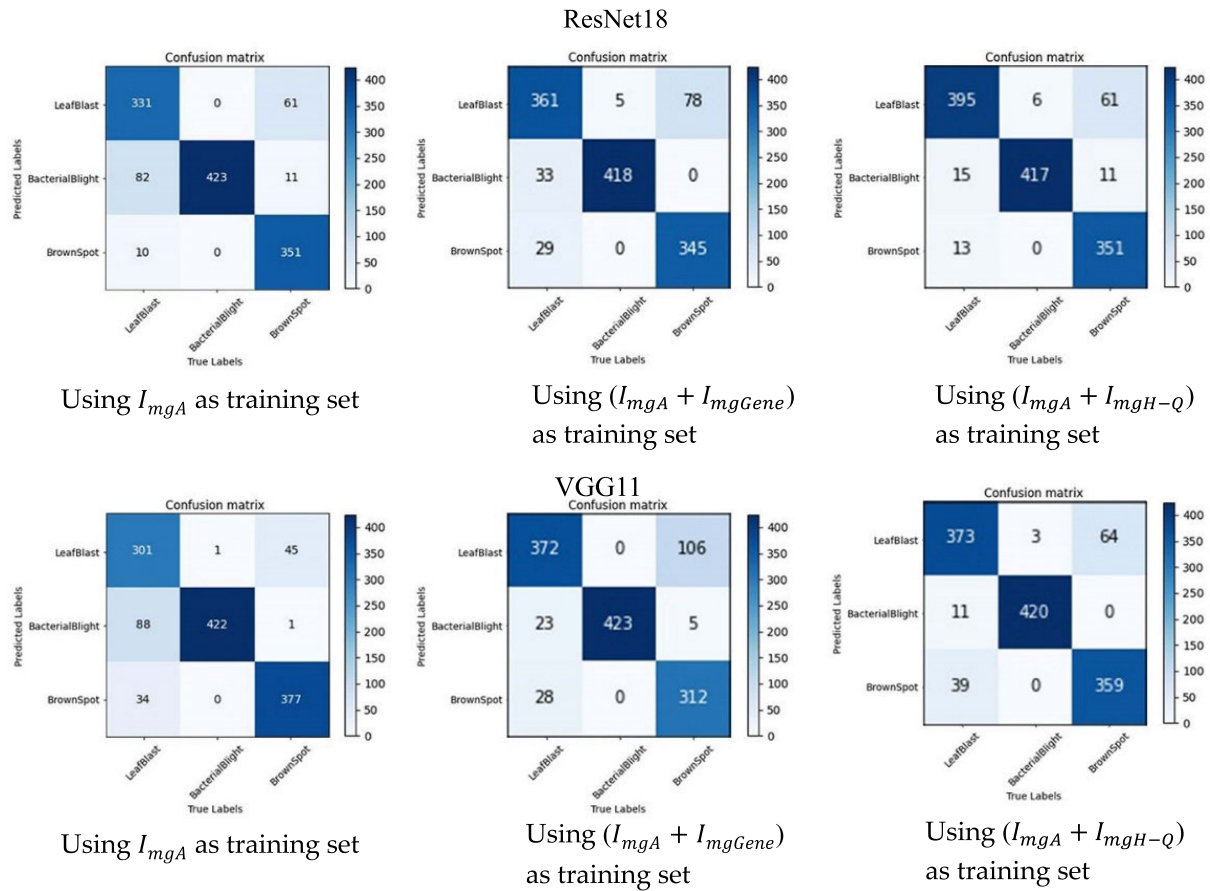
ResNet18

Using I_{mgA} as training set

Using $(I_{mgA} + I_{mgGene})$ as training set

Using $(I_{mgA} + I_{mgH-Q})$ as training set

VGG11

Using I_{mgA} as training set

Using $(I_{mgA} + I_{mgGene})$ as training set

Using $(I_{mgA} + I_{mgH-Q})$ as training set

FIGURE 12. **Confusion matrix on test set.**

respectively. When using the $I_{mgH-Q} + I_{mgA}$ dataset, the precision, recall, and F1-measure increased to 92.00, 91.67, and 91.83%, respectively, which were 4.13, 4.57, and 4.35% higher than those of the I_{mgA} dataset. Compared with the $I_{mgGene} + I_{mgA}$ dataset, these values were 3.27, 3.1, and 3.18% higher, respectively.

Table 7 shows the rice disease identification results of the best ResNet18 model. When VGG11 was used as the disease image recognition model, the accuracy using the I_{mgA} dataset was 86.68%, and the accuracy of the ResNet18 model using the $I_{mgGene} + I_{mgA}$ dataset was 87.23%. The accuracy of the dataset using $I_{mgH-Q} + I_{mgA}$ improved to 90.78%, which was 4.1% higher than that using the I_{mgA} dataset and 3.55% higher than that using the $I_{mgH-Q} + I_{mgA}$ dataset. The precision, recall, and F1-measure of the VGG11 model using the I_{mgA} dataset were 87.00, 86.70, and 86.84%, respectively. The precision, recall, and F1-measure of the VGG11 model using the $I_{mgGene} + I_{mgA}$ dataset were 87.80, 87.23, and 87.52%, respectively. When using the $I_{mgGene} + I_{mgA}$ dataset, the precision, recall, and F1-measure were 90.80, 90.80, and 90.80%, respectively, which were 3.80, 4.10, and 3.96% higher than those of the I_{mgA} dataset. Compared with the $I_{mgGene} + I_{mgA}$ dataset, these values were 3.00, 3.57, and 3.28% higher, respectively.

TABLE 7. **Rice disease recognition results of the best VGG11 model.**

Performance Metric	I_{mgA}	$I_{mgGene} + I_{mgA}$	$I_{mgH-Q} + I_{mgA}$
Accuracy	86.68%	87.23%	90.78%
Precision	87.00%	87.80%	90.80%
Recall	86.70%	87.23%	90.80%
F1-measure	86.84%	87.52%	90.80%

Based on these experimental results, RQ2 was answered. The enhanced data samples based on Opt-Real-ESRGAN further effectively improved the accuracy of image recognition.

The original disease dataset I_{mg} was divided into a training set I_{mgA} and test set I_{mgB} with a 1:1 ratio. We trained classification models using I_{mgA}, $I_{mgGene} + I_{mgA}$, and $I_{mgH-Q} + I_{mgA}$, respectively. Then we use I_{mgB} to test the performance of the classification models. Fig.12 shows the confusion matrix result on original test dataset I_{mgB}. The ResNet18 confusion matrix of the $I_{mgH-Q} + I_{mgA}$ dataset showed that the model had the best diversity. Compared with $I_{mgGene} + I_{mgA}$, the $I_{mgH-Q} + I_{mgA}$ dataset improved the identification rate of bacterial blight and brown spot. When $I_{mgGene} + I_{mgA}$ is used as a dataset training model, the recognition accuracy of one
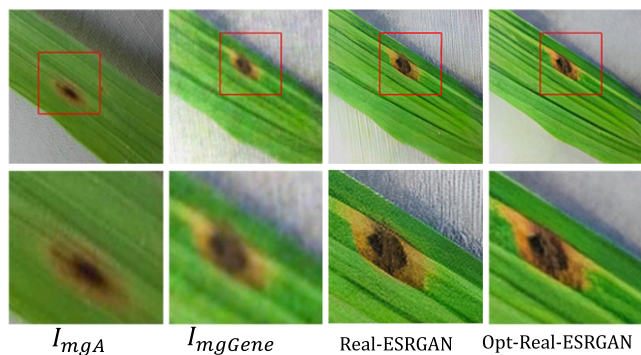
$$I_{mgA} \quad I_{mgGene} \quad \text{Real-ESRGAN} \quad \text{Opt-Real-ESRGAN}$$

FIGURE 13. Comparison of imaging effects at each stage.

TABLE 8. Comparison of Super-resolution models.

Super-resolution method	PSNR	SSIM	POS
Real-ESRGAN	25.31	0.78	3.82
Real-ESRGAN (Improved)	24.09	0.63	0.98

TABLE 9. Rice disease recognition results of the best ResNet18 model.

Performance Metric	I_{mgA}	I_{mgAH-Q}	$I_{mgH-Q}+I_{mgA}$
Accuracy	87.08%	87.86%	91.65%
Precision	87.87%	88.63%	92.00%
Recall	87.10%	87.86%	91.67%
F1-measure	87.48%	88.24%	91.83%

disease increases while that of another disease decreases. This is because the image quality of the super-resolution I_{mgH-Q} dataset was higher than that of I_{mgGene}. Higher image quality enabled the model to learn more disease features and further improved the accuracy of image recognition.

Compared with I_{mgA}, the overall disease recognition rate of the $I_{mgH-Q}+I_{mgA}$ dataset improved, and the overall diversity became very good. The diversity of the model became better, indicating that the more disease features the model learns, the better it can identify disease types and improve the overall generalization ability of the model. This is because the HQIA method generated high-quality rice leaf diseases; that is, the number of disease samples in the dataset expanded, and the quality of the expanded disease samples was higher than that of the traditional single-image generation network, further improving the accuracy of disease recognition and the overall diversity of the model.

3) COMPARISON BETWEEN OPT-REAL-ESRGAN AND REAL-ESRGAN

Opt-Real-ESRGAN was based on the improvement of Real-ESRGAN. This paper compared objective indicators and subjective effects.

Following mainstream methods, Real-ESGAN and Opt-Real-ESRGAN were used to upsample the I_{mgGene} dataset for X2 super-resolution enhancement, and the generated super-resolution result maps were evaluated for PSNR, SSIM, and POS. Two image super-resolution processing algorithms were compared and evaluated, and the results are shown in Table 8.

As shown in Table 8, the Real-ESGAN model and the Opt-Real-ESRGAN model had advantages and disadvantages. This is because the emphasis of the two models differed. Real-ESRGAN pursues more details of generation and obtains higher PSNR and SSIM indicators. Opt-Real-ESRGAN is faced with generated images containing more noise, focuses on the disease spots in super-resolution, and pursues faster generation speed. Fig. 13 shows the disease images generated by Real-ESRGAN and Opt-Real-ESRGAN. In comparison, both can achieve super-resolution enhancement of the rice disease image samples. However, as designed, Opt-Real-ESRGAN made some choices when enhancing the image

and only enhanced the leaf part, achieving a faster image generation speed. The details of super-resolution were only related to the effective features of disease image recognition. In contrast, although Real-ESRGAN had sharper detailed textures, it also enhanced the details of the background parts unrelated to disease recognition, which also led to its slow super-resolution generation speed. At the same time, these background features interfered with the performance of the disease image recognition model. Although Real-ESRGAN generated more details and textures, it also enhanced many background parts with texture details. The enhancement of these background textures not only seriously affected the speed of image generation but also produced a large number of noises features irrelevant to disease image recognition, which interfered with the performance of disease image models. Therefore, Opt-Real-ESRGAN was designed and selected for image super-resolution enhancement in HQIA in this paper.

4) INFLUENCE OF SUPER RESOLUTION TECHNOLOGY ON DISEASE RECOGNITION MODEL

According to RQ4, super-resolution enhancement is performed on I_{mgA} dataset to generate I_{mgAH-Q} dataset. The disease recognition model is trained using I_mgA dataset, I_{mgAH-Q} dataset and $I_{mgH-Q}+I_{mgA}$ dataset respectively. The above model is then tested using the ImgB dataset. Verify and answer RQ4, how does the disease recognition accuracy change by only using super-resolution enhancement without augmenting the dataset?

Table 9 and Table 10 show the best results for rice disease identification using only the super resolution model. In the experimental results, we can analyze that the performance of disease recognition model with super resolution I_{mgAH-Q} data set is improved to a certain extent compared with the original data set I_{mgA}, but the performance is weaker than that of $I_{mgH-Q}+I_{mgA}$ data set. This is because the image

TABLE 10. Rice disease recognition results of the best VGG11 model.

Performance Metric	I_{mgA}	I_{mgAH-Q}	$I_{mgH-Q}+I_{mgA}$
Accuracy	86.68%	88.89%	90.78%
Precision	87.00%	89.20%	90.80%
Recall	86.70%	88.86%	90.80%
F1-measure	86.84%	89.03%	90.80%

super resolution technology makes the image clear and sharp, which makes it easier for the disease recognition model to learn the hidden features of the disease, but does not make the image produce more hidden features. However, the HQIA method in this paper uses WGAN-GP for image augmentation. The expanded new image data set provides more hidden features for the model for training, improves the overall training space of the model and makes the model have better generalization ability. HQIA also combines the advantages of super resolution to clearly sharpen the image data set augmented by WGAN-GP, making its disease hiding features easier to learn. Image super resolution is a full use of the existing data set, while WGAN-GP is an extension of the data set. Only the combination of the two can give full play to the best performance, which also verifies the rationality of HQIA method.

V. CONCLUSION

Rice is a staple food crop in the world, and disease is one of the most important factors affecting rice yield. It is of great significance to detect and determine the type of disease in time for disease prevention and control. Relying on deep learning for rice disease recognition is the current mainstream direction. Deep learning models need sufficient training samples to support them in the training process; otherwise, overfitting occurs, resulting in model failure. However, in the field of agriculture, there are common problems, such as difficulty in obtaining high-quality disease samples and high cost. To solve these problems, this paper proposed a dual GAN network for a high-quality rice leaf disease image data augmentation HQIA method. This method used a WGAN-GP network to generate image samples and then an Opt-Real-ESRGAN network to perform super-resolution enhancement on the generated image samples.

The experimental results showed that the HQIA method in this paper improved recognition accuracy by 4.57% using ResNet18 and 4.1% using VGG11 compared to without data augmentation. Compared with the traditional WGAN-GP data augmentation method, the accuracy improved by 3.08% using ResNet18 and 3.55% using VGG11.

The HQIA method proposed in this paper can also be extended to other plant disease identification, and we hope to reduce the impact on researchers due to insufficient data sets. The HQIA method proposed in this paper also has some shortcomings, and we will make improvements in the following two aspects in the future.

1) Currently, we only study the image data augmentation method based on simple background. In the future, we will continue to study the data augmentation method based on complex background under natural conditions.

2) We improved the Real-ESRGAN network to reduce the consumption of overall computing performance, but the overall computing performance is still relatively high consumption. In the future, we should aim to reduce the overall algorithm complexity and realize the lightweight deployment of the overall algorithm.

REFERENCES

[1] D. F. Zhu, "Analysis of status and constraints of Rice production in the world," *Scientia Agricultura Sinica*, vol. 43, no. 3, pp. 474–479, 2010.

[2] L. B. Liu and G. M. Zhou, "Identification method of Rice leaf blast using multilayer perception neural network," *Trans. CSAE*, vol. 25, no. 2, pp. 213–217, 2009.

[3] Z. X. Guan, "Study on recognition method of Rice disease based on image," *Chin. J. Rice Sci.*, vol. 24, no. 5, pp. 497–502, 2010.

[4] T. T. Liu, T. Wang, and L. Hu, "Rhizocotonia solani recognition algorithm based on convolutional neural network," *Chin. J. Rice Sci.*, vol. 33, no. 1, pp. 90–94, 2019.

[5] T. Liu, "Recognition of Rice leaf diseases based on computer vision," *Scientia Agricultura Sinica*, vol. 47, no. 4, pp. 664–674, 2014.

[6] C. Charliepaul, "Classification of Rice plant leaf using feature matching," *Int. J. Eng. Technol. Sci.*, vol. 1, no. 2, pp. 290–295, 2014.

[7] G. Anthonys and N. Wickramarachchi, "An image recognition system for crop disease identification of paddy fields in Sri Lanka," in *Proc. Int. Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2009, pp. 403–407.

[8] S. P. Huang, C. Sun, Q. Long, X. Ma, and W. J. Wang, "Rice panicle blast identification method based on deep convolution neural network," *Trans. Chin. Soc. Agricult. Eng. Trans. CSAE*, vol. 33, no. 20, pp. 169–176, 2017.

[9] Y. L. Tan, "Image recognition of Rice diseases based on deep convolutional neural network," *J. Jing Gang Shan Univ. Natural Sci.*, vol. 40, no. 2, pp. 38–45, 2019.

[10] S. Bhattacharya, A. Mukherjee, and S. Phadikar, "A deep learning approach for the classification of Rice leaf diseases," *Intell. Enabled Res.*, vol. 1109, pp. 61–69, Jan. 2020.

[11] Q. Xu, M. Zhang, Z. Gu, and G. Pan, "Overfitting remedy by sparsifying regularization on fully-connected layers of CNNs," *Neurocomputing*, vol. 328, pp. 69–74, Feb. 2019.

[12] S. S. Du et al., "How many samples are needed to learn a convolutional neural network?" in *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Montreal, QC, Canada, Dec. 2018, pp. 371–381.

[13] L. Y. Mo, "Research on personalized recommendation system based on linked data," M.S. thesis, South China Univ. Technol., Guangdong Province, Guangzhou, China, 2018.

[14] J. Sun, W. J. Tan, H. P. Mao, X. H. Wu, Y. Chen, and L. Wang, "Recognition of multiple plant leaf diseases based on improved convolutional neural network," *Trans. Chin. Soc. Agricult. Eng. Trans. CSAE*, vol. 33, no. 19, pp. 209–215, 2017.

[15] J. C. Ma, K. M. Du, F. X. Zheng, L. X. Zhang, and Z. F. Sun, "Disease recognition system for greenhouse cucumbers based on deep convolutional neural network," *Trans. Chin. Soc. Agricult. Eng. Trans. CSAE*, vol. 34, no. 12, pp. 186–192, 2018.

[16] W. X. Bao, "Apple leaf disease recognition based on improved convolutional neural network," *J. Anhui Univ. Natural Sci. Ed.*, vol. 45, no. 1, pp. 53–59, 2021.

[17] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Conf. Adv. Neural Inf. Process. Syst.* Montreal, QC, Canada, 2014, pp. 2672–2680.

[18] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "Characterization and computation of local Nash equilibria in continuous games," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, 2013, pp. 917–924.

[19] G. Hu, H. Wu, Y. Zhang, and M. Wan, "A low shot learning method for tea leaf's disease identification," *Comput. Electron. Agricult.*, vol. 163, Aug. 2019, Art. no. 104852.

[20] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[21] Q. Wu, Y. Chen, and J. Meng, "DCGAN based data augmentation for tomato leaf disease identification," *IEEE Access*, vol. 8, pp. 98716–98728, 2020.

[22] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.

[23] J. J. Bird, "Fruit quality and defect image classification with conditional GAN data augmentation," *Scientia Horticulturae*, vol. 293, Feb. 2022, Art. no. 110684.

[24] J. J. Liang, J. J. Wei, and Z. F. Jiang, "Generative adversarial networks GAN overview," *J. Frontiers Comput. Sci. Technol.*, vol. 14, no. 1, pp. 1–17, 2020.

[25] J. H. Luo and J. X. Wu, "A survey on fine-grained image categorization using deep convolutional features," *Acta Automatica Sinica*, vol. 43, no. 8, pp. 1306–1318, 2017.

[26] M. Z. Zeng, "Research on crop disease image data augmentation method based on generative adversarial network," M.S. thesis, Pattern Recognit. Intell. Syst., Univ. Sci. Technol. China, Anhui, China, 2021.

[27] R. C. Sun, *Research on Generating Data Enhancement Technology of Networks in Plant Disease Classification*. Chongqing China: Chongqing Univ., 2021.

[28] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1905–1914.

[29] Y. Gao, Z. Liu, P. L. Qin, and L. F. Wang, "Medical image super-resolution algorithm based on deep residual generative adversarial network," *J. Comput. Appl.*, vol. 2018, vol. 38, no. 9, pp. 2689–2695.

[30] A. Li and X. Q. Song, "Remote sensing image super-resolution reconstruction based on generative adversarial networks," *Opt. Optoelectron. Technol.*, vol. 17, no. 6, pp. 39–44, 2019.

[31] I. Gulrajani, "Improved training of Wasserstein GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, p. 30.

[32] H. Xu, "Lipschitz constrained parameter initialization for deep transformers," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, Jul. 2020, pp. 397–402.

[33] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based Rice leaf disease identification using support vector machine," *Comput. Electron. Agricult.*, vol. 175, Aug. 2020, Art. no. 105527.

[34] *Hefei Institute of Intelligent Machinery, Chinese Academy of Sciences. Rice False Smut Dataset [DS/OL]*, Science Data Bank, Beijing, China, 2019, doi: 10.11922/sciencedb.p00001.00002.

[35] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 126–135.

[36] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 114–125.

[37] X. Wang, K. Yu, and C. Dong, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 606–615.

ZHAO ZHANG received the B.S. degree in electronic information science and technology from the Yantai Institute of Technology, China, in 2019. He is currently pursuing the M.S. degree in agricultural engineering and information technology with Yunnan Agricultural University, Kunming, China. His current research interests include machine learning, deep learning, generate countermeasure networks, image super-resolution, and data augmentation.

QUAN GAO received the M.S. degree from Yunnan University, China, in 2009. He is currently an Associate Professor with the College of Big Data, Yunnan Agriculture University, Kunming, China. His research interest includes agricultural engineering.
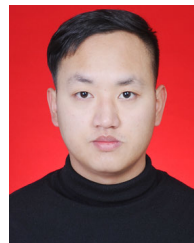
LIRONG LIU was born in Luliang, Shanxi, China, in 1997. He received the bachelor's degree in software engineering from Shanxi Agricultural University, in 2020. He is currently pursuing the master's degree in agricultural engineering and information technology with Yunnan Agricultural University. His research interest includes agricultural engineering.

YUN HE received the B.S., M.S., and Ph.D. degrees from Yunnan University, China, in 2011, 2015, and 2019, respectively. He is currently a Lecturer with the College of Big Data, Yunnan Agriculture University, Kunming, China. His research interests include agricultural engineering, software engineering, and machine learning.

• • •