

# Aerial Image Quality Enhancement Using Enhanced Super-Resolution Generative Adversarial Networks for Rice Plant Disease Identification

Dwi Ratna Sulistyaningrum, Al Fatoni Nugroho Putra, Budi Setiyono  
and Muhammad Khusni Nailul Mubarak

Department of Mathematics  
Institut Teknologi Sepuluh Nopember  
60111, Surabaya, Indonesia

dwiratna@its.ac.id; alfatoni7i@gmail.com; masbudisetiyono@gmail.com; nallsst16@gmail.com

Received December 26, 2024, revised March 11, 2025, accepted March 11, 2025.

---

**ABSTRACT.** *Disease identification in rice plants is one of the essential elements in maintaining agricultural productivity in Indonesia. Tungro and blast are two diseases that often attack rice plants in Indonesia. Aerial imagery captured with the help of drones is a promising alternative for monitoring diseased areas on farms more efficiently. However, aerial images produce low resolution when zoomed in, limiting the ability to identify diseases. The enhanced super-resolution generative adversarial networks (ESRGAN) method can increase the resolution of aerial images so that the classification of rice plant diseases with the convolutional neural networks (CNN) classification method performs better. The test results show that identifying rice plant diseases with aerial image data has an accuracy of 86.96% when the image data is enhanced in resolution with ESRGAN, and its accuracy is 79.97% without ESRGAN.*

**Keywords:** Rice Plant Diseases, Aerial Image, Enhanced Super Resolution Generative Adversarial Networks.

---

**1. Introduction.** In Indonesian, the agricultural sector plays a vital role in the country's economic system. The conditions of natural resources and climate in Indonesia make the agricultural sector from the past to the present still experience development over time. According to the Central Bureau of Statistics, in the first quarter of 2023, the agricultural sector was called the most dominant sector, with a growth rate of 0.34% and a contribution of 11.77% to the gross domestic product (GDP). Then, according to the International Monetary Fund (IMF), Indonesia is among the 20 countries with the largest economies in the world, which was recorded in the October 2023 edition of the World Economic Outlook database. This achievement cannot be separated from the role of farmers and the government, since it was still done manually using conventional methods; until now, it has developed sophisticated supporting tools.

The progress of the agricultural sector in Indonesia is obtained through the role of innovation of agricultural managers, which is increasingly leading to the digitization of agriculture to obtain accurate data about problems that have been faced in the past, especially the problem of pests or diseases that attack rice plants. These problems will directly impact rice plants, increasing the frequency of crop failures and impacting the progress of the agricultural sector in Indonesia. Various types of diseases in rice plants that often occur in Indonesia include blast disease, tungro disease, leaf blight, and pests

such as leafhoppers, locusts, and rats [1]. The season in Indonesia from 2023 to early 2024 is very unpredictable, making it difficult for farmers to deal with several diseases and pests quickly in rice fields. The limited ability of humans, including farmers, to monitor rice fields makes diagnosing and handling a disease in rice very late. It causes rice plants to be damaged to cause crop failure [2]. Therefore, the development of innovations in tools to identify rice diseases is needed as soon as possible so that agricultural managers can diagnose the type of rice disease and ensure that the proper treatment is carried out quickly.

The characteristics of rice plants include leaves that are pretty small and elongated, then planted nearby, allowing the threat of diseases and pests to be straightforward and fast. The identification of diseases in rice plants in Indonesia today is still reasonably conventional, namely by looking closely at each rice plant, which, of course, requires a long time and sufficient insight so that handling can be done, even though if observed thoroughly with the naked eye from a distance it is clear that the rice plant is indicated to be affected by a disease or pest. In the current development in the field of technology, there is a field of programming that is used to help identify a disease with data in the form of images (pictures or videos), namely image processing technology [3]. The development of image processing methods has now been combined with Artificial Intelligence technology. However, most previous studies still use the object of the leaf image or the object is taken from a close distance, so there is no difference from the conventional way. So, it is necessary to use image data from a certain distance to see the expanse of rice plants so that it can accurately and quickly determine the location of diseased rice areas accurately and quickly. A sophisticated Unmanned Aerial Vehicle (UAV) tool has been developed, one example of which is the drone [4]. This multispectral drone technology is specifically used in the agricultural sector, especially to mapping detailed and accurate land conditions, crop conditions, and pests and diseases that exist in it. These drones are only used in a few areas with sufficient funds, while many areas in Indonesia still need more funds to buy these multispectral drones. Another drone technology, RGB drones, is one solution for areas that need help to afford multispectral drones. With the help of RGB drones, it is possible to get diseased areas in a large expanse of rice fields. However, the results of RGB drone capture at a certain distance will produce objects in the image to be less evident or blurred, unlike multispectral drones that have been equipped with sophisticated tools, so it is necessary to use a technique to help image processing on the RGB drone capture to see objects. Not only with Artificial Intelligence but currently, there is also a lot of research on Super-Resolution on images that transform images and videos that initially had low resolution into high resolution so that the images produced by drones can be clarified and can be used for better object identification [5].

Based on the explanation above, it can be seen that the characteristics of rice need super-resolution techniques to get rice leaf objects that are visible for disease. Therefore, in this research, Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) is applied to improve the quality of aerial imagery in identifying plant diseases, especially in rice plants [6]. In previous studies, the ESRGAN method has achieved good results in helping disease classification. However, it is only on a few leaves and taken from a reasonably close distance, and the leaves that are used as objects are relatively broad in contrast to rice leaves, which have a relatively small and slender leaf size [7]. Furthermore, aerial image data was obtained from a camera mounted on a drone. The ESRGAN super-resolution technique is used to obtain aerial images of diseased areas with higher resolution to determine the type of disease suffered by the rice plant. Convolutional Neural Network (CNN) is a deep learning method used in various studies to recognize objects [8, 9, 10, 11, 12]. This study uses the Convolutional Neural Network (CNN) method to

identify rice plant diseases with the EfficientNet architecture. Processing images generated by drones using super-resolution techniques using ESRGAN and supporting methods for disease identification using CNN used in this study are expected to help the agricultural sector in Indonesia.

The research focus can be observed through the following questions: First, how to apply Enhanced Super-Resolution Generative Adversarial Networks for aerial image quality enhancement in plant disease identification. Second, what are the performance results of the Enhanced Super-Resolution Generative Adversarial Networks method to improve the quality of aerial imagery in assisting plant disease identification? There are also some limitations in this research: First, the research object is rice plants with blast and tungro diseases. Second, the aerial image data is taken during sunny conditions. Third, one aerial image contains only one type of disease. The rest of this paper is organized as follows: Section 2 introduces related works related to image quality improvement, especially in aerial imagery. Section 3 shows the workflow from the aerial image quality enhancement system to identifying diseases in rice. The experimental results are presented and discussed in Section 4. The last section, Section 5, is dedicated to conclusions and suggestions for future work.

**2. Related Work.** Research on super-resolution techniques in the agricultural sector has been conducted in recent years. In a study [13], the super-resolution convolutional neural network (SRCNN) method was successfully applied to improve the image quality of leaves to recover the details. The leaf image object used is a single tomato leaf image that has a disease on the leaf. In addition to using SRCNN, the study also conducted a disease classification process using high-resolution, low-resolution, and super-resolution image data to evaluate the effectiveness of the super-resolution method in disease classification. AlexNet is the architecture used in the classification process.

Moghimi et al. [14] examined hyperspectral imagery for yield phenotyping in wheat. Hyperspectral cameras are mounted on unmanned aerial vehicles (UAVs) to collect aerial imagery with high spatial and spectral resolution. The apparent advantage of imagery with hyperspectral channels is the ability to capture yield variation at the subplot scale. Thus, the research resulted in a performance that can facilitate the phenotyping process of high-throughput results.

Agricultural research continues to develop again with varied objects. Selvaraj et al. [15] have written a research paper on the detection and classification of banana plants and provide more information about banana plants, especially on the health status of these plants. This research combines images obtained from the air using unmanned aerial vehicles (UAVs) and satellites with artificial intelligence technology in the form of machine learning.

Albattah et al. [11] have successfully identified diseases with a class of 32 types of diseases and are in the test phase using drone-generated images. The method used in that study is excellent, so it gets a perfect accuracy rate even though the leaf object is quite large. However, this research only targets disease classification in aerial image data without segmentation of diseased areas and additional super-resolution.

**3. Aerial Image Quality Enhancement System Framework for Disease Identification.** The design of the working stages of the system regarding the use of Enhanced Super-Resolution Generative Adversarial Networks to improve the identification of plant diseases in aerial images is shown in Figure 1. The proposed system consists of 4 main processes: (i) data preparation, (ii) aerial imagery segmentation, (iii) super-resolution model building, (iv) disease classification modeling. In this study, the image segmentation

process uses the Visible Atmospherically Resistant Index (VARI) method. The super-resolution process uses the Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) model. Classification of plant diseases with the Convolution Neural Network (CNN) method. Detailed descriptions of each process are explained in the following subsections.

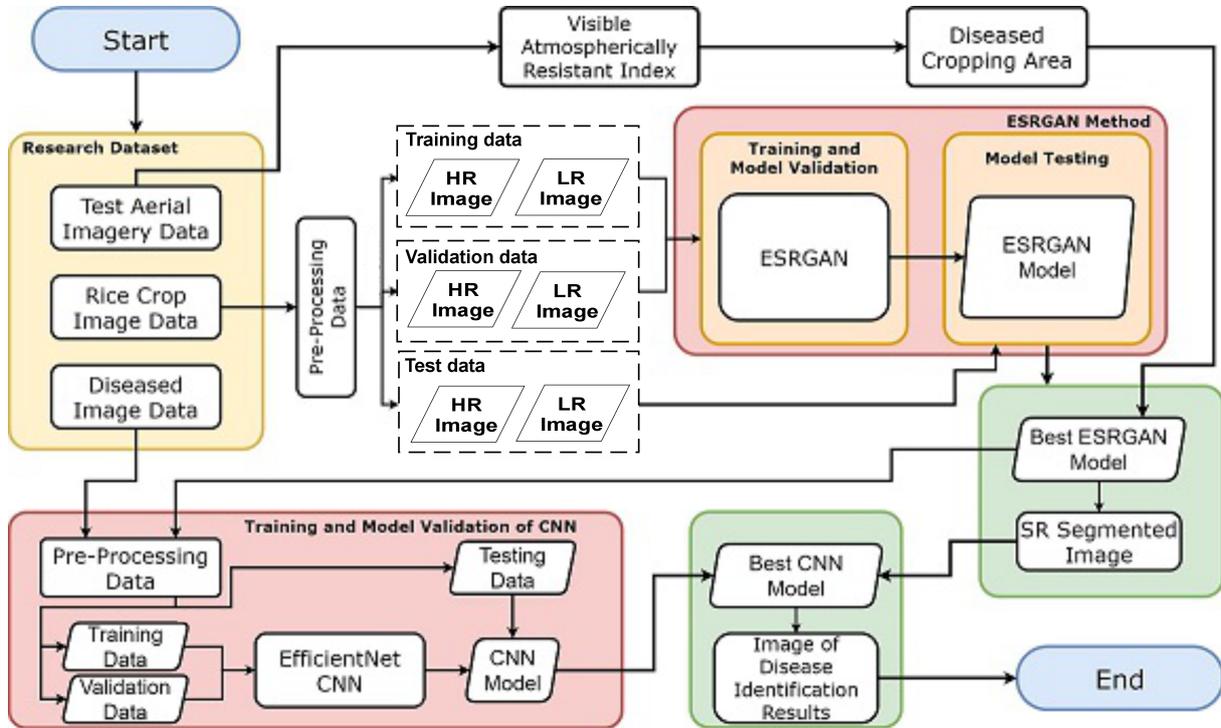


FIGURE 1. System work stages diagram

**3.1. Preparing Data.** Data preparation is a stage to prepare the dataset used to identify diseases in aerial imagery. Aerial image data is primarily taken using a DJI Mini 3 brand drone camera. This data comes from various locations in Tulungagung district, East Java. The collection of aerial imagery data focuses on the rice field area in the centre so that it can focus on the expanse of rice plants only and is carried out during sunny conditions. The illustration of the data collection used in this study is shown in Figure 2.



FIGURE 2. Simulation of research data collection

The aerial image data in this research is taken from the extraction of video frames generated according to the illustration above. The result of this video makes it possible to get more images in a short time to minimize the use of batteries in the drone and make it more efficient. There is also another advantage of selecting frames that allows

production images to focus on the area where the disease is present. The video data that was successfully obtained has an average duration for each video of 60 seconds with MPEG-4 Part 14 (.MP4) format with 4K resolution or  $3840 \times 2160$  pixels and has a framerate of 30 FPS (Frame per Second). When recording this aerial video, the drone's height is about  $\pm 10$  meters measured from the top of the rice plant and uses an angle of  $90^\circ$ . Then, the aerial image data is broken down into three types of datasets.

1. *Test Aerial Imagery Data.* The test aerial image data is only generated from video frame extraction without processing. The size is  $3840 \times 2160$  pixels, and the image is taken if there are diseased areas so that the area can be segmented later. The total of the test aerial image data is ten images shown in Table 1.

TABLE 1. Test Aerial Imagery Data

No	Disease	Image Name
1	Blast	Blast_10m_1_4K.jpg
2		Blast_10m_2_4K.jpg
3		Blast_10m_3_4K.jpg
4		Blast_10m_4_4K.jpg
5		Blast_10m_5_4K.jpg
6	Tungro	Tungro_10m_1_4K.jpg
7		Tungro_10m_2_4K.jpg
8		Tungro_10m_3_4K.jpg
9		Tungro_10m_4_4K.jpg
10		Tungro_10m_5_4K.jpg

2. *Rice Crop Image Data.* The image data of rice plants is obtained from the extraction of aerial video frames. The object magnification process is carried out on the image to become the ground truth of the ESRGAN method. This rice plant image data aims to train ESRGAN to recognize and create SR images of rice plants. Then, the process produced 60 images of rice plant frames. From the 60 images of rice plant frames, the sub-image extraction process is carried out into a square image with a size of  $512 \times 512$  pixels and a total of 28 sub-images for one frame image. So, the total image data of rice plants is 1,680 images. Figure 3 shows some images resulting from the sub-image extraction.

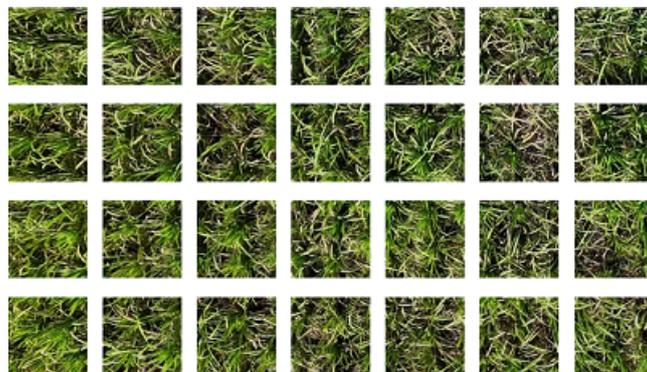


FIGURE 3. Sample of sub-image extraction results

3. *Diseased Image Data.* The diseased image data was obtained from the extraction of aerial video frames, and a manual cropping process was carried out on diseased areas only according to the farmers' statements in the area and validated by the Tulungagung Regency Agriculture Office. From the results of this process, ten frame images of different sizes were taken for each disease used in the study, namely blast and tungro, so the total image obtained was 20 diseased frame images. From the 20 diseased frame images, the sub-image extraction process is carried out by breaking the original image into  $128 \times 128$  pixels so that the total diseased image is 600, with each disease being 300. The results of the diseased sub-image extraction are visualized in Figure 4.



FIGURE 4. Sample of the diseased sub-image extraction result

3.2. **Aerial Imagery Segmentation Process.** Image segmentation generally aims to separate objects from the background in an image based on differences in brightness or darkness. In this research, segmentation aims to identify the area of diseased rice, assuming that diseased rice has a different colour from healthy rice. The image segmentation process in this research uses the Visible Atmospherically Resistant Index (VARI) method. Then, the cropping process is carried out on the segmented part to focus more on knowing the type of disease [16], especially blast and tungro suffered by the rice plant. The VARI segmentation method has the following equation:

$$\text{VARI} = \frac{R_{green} - R_{red}}{R_{green} + R_{red} - R_{blue}} \quad (1)$$

where  $R$  denotes the reflectance level of the wavelengths in the three-channel images: red, green, and blue.

The matrix resulting from the VARI calculation is then subjected to thresholding. Thresholding works by adjusting the number of pixel degrees in the image and produces an output in the form of a binary image (black and white), which is mathematically written:

$$g(x, y) = \begin{cases} 0 & , f(x, y) \geq T \\ 1 & , f(x, y) < T \end{cases} \quad (2)$$

with  $f(x, y)$  is the pixel value or degree that has been generated from the VARI calculation. Meanwhile,  $g(x, y)$  is the value after thresholding segmentation, and the  $T$  value is the limit value of thresholding. The value of 1 produces white pixels, while 0 produces black pixels. In the case of this research, because each image has different conditions, the  $T$  value is determined by finding the 65th percentile value of the elements in the VARI calculation matrix. An illustration of the VARI to thresholding process is shown in Figure 5.

The thresholding results in the illustration above will be cropped on pixels that have a value of 1, which indicates that the area is affected by a disease. These areas are cropped to produce an image focused on the diseased area only and stored for the next stage. However, in the cropped image, there are still sub-areas unaffected by the disease, so

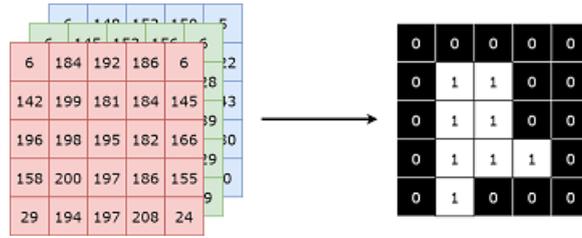


FIGURE 5. Illustration of image segmentation process with VARI

a sub-area extraction process is required by dividing the image into sub-images with a size of  $128 \times 128$  pixels. Besides focusing on the diseased area, this sub-image extraction process also aims to be the input image into the ESRGAN.

**3.3. Super Resolution Model Building.** Super-resolution modelling is the essential process in this research. This process requires data obtained from rice plant image data to be entered into the Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) method. The GANs method is a neural network model that uses two main parts, the generator and the discriminator, that work in an adversarial manner to achieve specific goals. The ESRGAN generator architecture uses 23 residuals on the residual dense block (RRDB), convolutional layer, Leaky Rectified Linear Unit (LReLU), scaling, and upsampling layer [17]. All layers have a kernel size of  $3 \times 3$  and stride shift size = 1. Meanwhile, the ESRGAN discriminator uses seven convolutional layers: batch normalization and LReLU block. The structure of the ESRGAN generator network is shown in Figure 6, while the structure of the ESRGAN discriminator network is shown in Figure 7 [17].

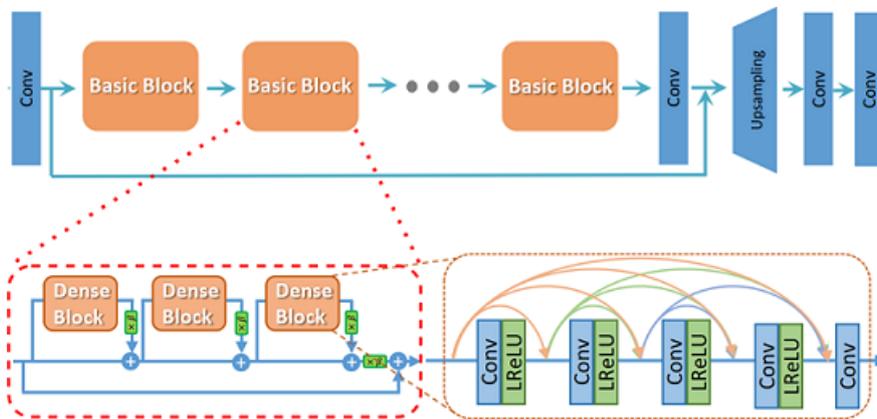


FIGURE 6. ESRGAN generator network



FIGURE 7. ESRGAN discriminator network

At this stage, the model training process is carried out on the Enhanced Super Resolution Generative Adversarial Network (ESRGAN) method, which starts by entering the LR image dataset resulting from the downgrade of 4x aerial images of rice plants into the generator network, which then the network will produce a new image or what is commonly called a fake image (in the form of an SR image) which is similar to the original. Then, the SR and HR image datasets are forwarded to the discriminator. The discriminator determines whether the generator-generated image looks realistic like the original image with the HR image as the ground truth [18]. The distribution of training, validation, and testing data is 80%, 10%, and 10% respectively. After that, we continued to look at the model results using assessment metrics in PSNR and SSIM [19].

$$\text{PSNR} = 20 \log_{10} \left( \frac{255}{\sqrt{\text{MSE}}} \right) \quad (3)$$

$$\text{SSIM}(X, Y) = \left( \frac{2(\mu_X \mu_Y) + C_1}{\mu_X^2 + \mu_Y^2 + C_1} \right) \left( \frac{2\sigma_{XY} + C_2}{\mu_X^2 + \mu_Y^2 + C_2} \right) \quad (4)$$

Relativistic average discriminator tries to predict the probability that the real image  $x_r$  is relatively more realistic than the fake one  $x_f$ . RaD is denoted by  $D_{Ra}$  which is formulated:

$$D_{Ra}(x_r, x_f) = \sigma (C(x_r) - \mathbb{E}_{x_f}[C(x_f)]) \quad (5)$$

where  $\mathbb{E}_{x_f}[\cdot]$  represents the average taking operation for all false data and  $\sigma$  is a sigmoid function and  $C(x)$  is the untransformed discriminator output. Then, discriminator loss is formulated as:

$$L_D^{Ra} = -\mathbb{E}_{x_r} [\log (D_{Ra}(x_r, x_f))] - \mathbb{E}_{x_f} [\log (1 - D_{Ra}(x_f, x_r))] \quad (6)$$

and adversarial loss for generator in symmetric form as in equation:

$$L_G^{Ra} = -\mathbb{E}_{x_r} [\log (1 - D_{Ra}(x_r, x_f))] - \mathbb{E}_{x_f} [\log (D_{Ra}(x_f, x_r))] \quad (7)$$

with  $x_f = G(x_i)$  and  $x_i$  is the image of input. Since the adversarial loss for generator contains both  $x_r$  and  $x_f$ , so the generator in ESRGAN benefits from both the gradient of generated data and real data in *adversarial training*, whereas in SRGAN only the generated part is applicable. This modification of the discriminator in ESRGAN helps learn sharper edges and more detailed textures.

**3.4. Disease Classification Modeling.** A disease classification model is developed to identify diseases in rice plants. This process requires diseased aerial imagery data as input for the Convolution Neural Network (CNN) training process with the EfficientNet-B3 architecture. Figure 8 shows the EfficientNet architecture used in this study [12].

Data pre-processing, which includes data augmentation techniques, is carried out in this research. The data augmentation techniques used are translation, rotation, reflection, dilation, and brightness. Then, the system assigns a disease label to the pre-processed diseased aerial image data based on the ground truth and later inputs it into the layers of the EfficientNet CNN architecture [20]. There are three main layers: ConvReLU Layer, MBConv Layer, and Pooling Layer. Each layer is used for the feature extraction process to recognize diseases in rice plants and ends with the classification process [21]. In addition to performing data augmentation, the system downgrades disease image data to LR images and carries out the ESRGAN process using a model created during ESRGAN training. This research designs the image data as one of the scenarios, so the CNN model consists of two main models: the model resulting from LR image training of rice disease and the model resulting from SR image training of rice disease. This training process includes two disease classes: blast and tungro. Then, the system splits the three disease datasets

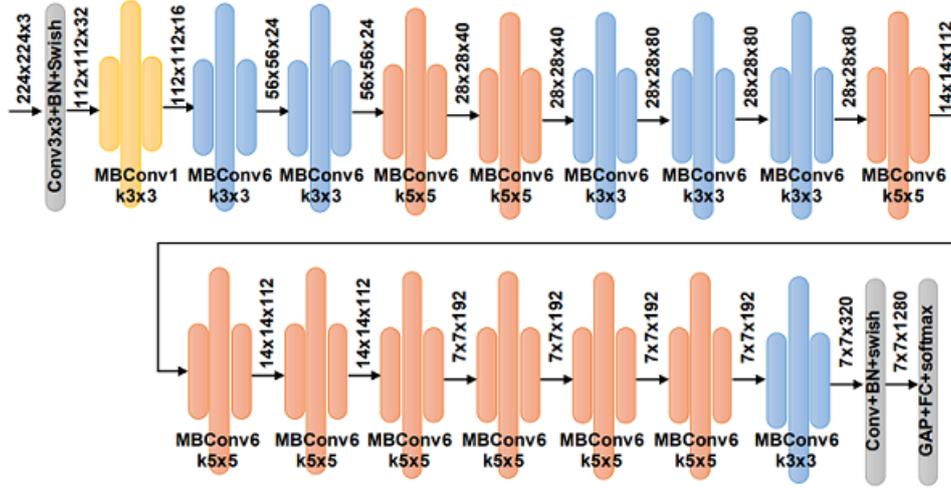


FIGURE 8. EfficientNet architecture

into 80% for training data, 10% for validation data, and 10% for test data for each disease. Afterwards, it evaluates the model performance using accuracy, precision, and recall metrics [22].

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

**4. Experiment Result and Discussion.** This section describes the experimental results of each process stage and evaluates the performance results of the method used. Model testing uses predetermined test data. The results of the experimental process include the image segmentation process of diseased rice areas, the super-resolution process using ESRGAN and the disease classification process using the CNN *EfficientNet-B3* architecture.

**4.1. Aerial Image Segmentation Process Test.** At this stage, the system performs the segmentation process on aerial images of rice plants. Figure 9(a) visualizes examples of aerial images with Blast disease obtained and used in this research trial. Then, the system segments the diseased area using the VARI method, applying the calculations described in the previous section. The results of the VARI segmentation are shown in Figure 9(b). When converted into binary images, the VARI results still contain some small noise areas. These small areas increase the number of processed images, so the system first reduces noise using connected components before cropping. Figure 9(c) shows the final result of the segmentation process using VARI.

Based on the final results of aerial image segmentation of diseased areas, the cropping process was carried out on each diseased area and eight diseased areas were obtained. Figure 10 shows the cropped sub-images.

In the following process, the sub-image extraction of diseased areas with a sub-image size of  $128 \times 128$ . Eight segmentation areas are obtained; Figure 11 visually shows the results of sub-image extraction in segmentation area 7.

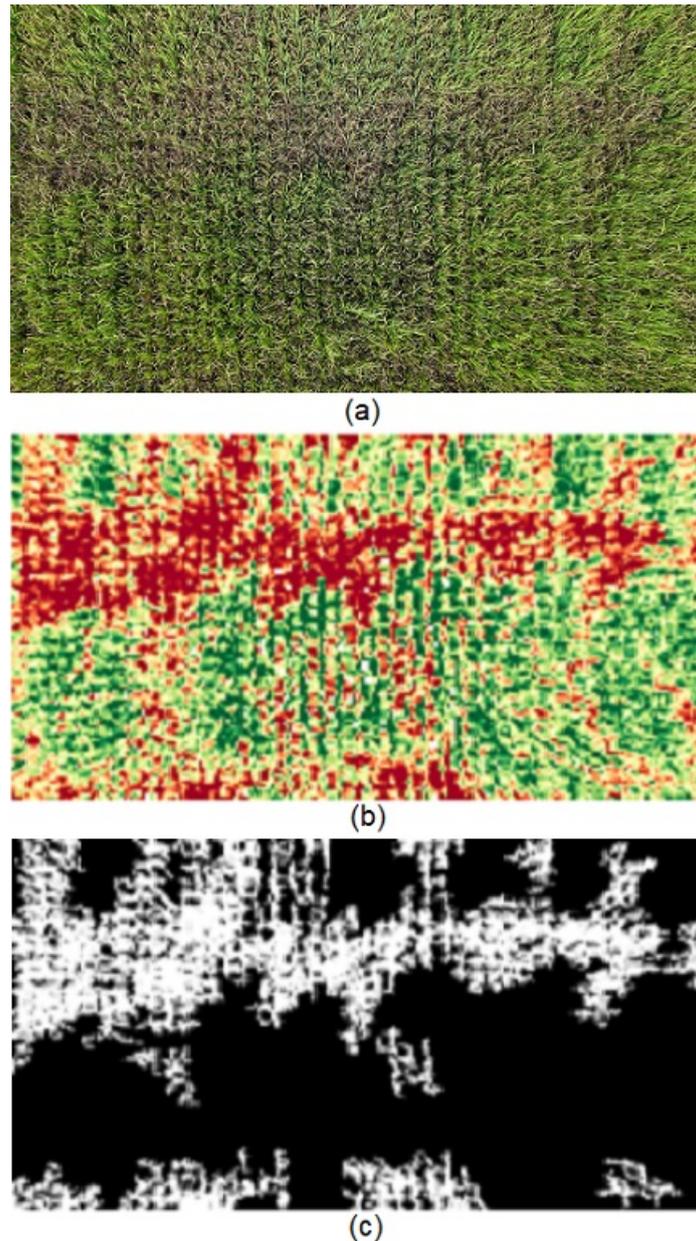


FIGURE 9. (a)Sample Blast Image (b)After the VARI Process (c)Final Result of Segmentation Diseased Area

With the same segmentation process on each test image, the following Table 2 presents the image name and the number of areas obtained due to VARI and cropping segmentation.

**4.2. Aerial Imagery Super Resolution Process Test.** There are 30 4K images, and then the sub-image extraction process is carried out, where the original image is split into 28 small images with a size of  $512 \times 512$ . The data splitting divides the dataset into three parts: training data totalling 1344 images, validation data totalling 168 images, and test data totalling 168 images. Training data is used to train the model. The validation data provides an unbiased evaluation of the model fit on the training data when tuning the model's hyperparameters. The test data provides an unbiased evaluation of the final model that fits the training data. The sub-image extraction results in the three datasets

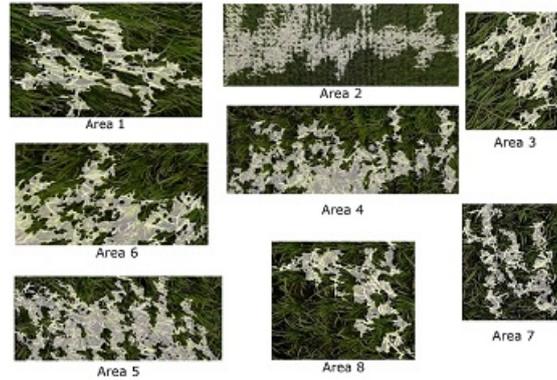


FIGURE 10. Diseased Area Cropping Results



FIGURE 11. Results of sub-image extraction

TABLE 2. Cropping Result of Diseased Area on Aerial Image

Disease	Image Name	Diseased Areas
Blast	Blast_10m_1_4K.jpg	8
	Blast_10m_2_4K.jpg	8
	Blast_10m_3_4K.jpg	7
	Blast_10m_4_4K.jpg	8
	Blast_10m_5_4K.jpg	5
Tungro	Tungro_10m_1_4K.jpg	4
	Tungro_10m_2_4K.jpg	9
	Tungro_10m_3_4K.jpg	5
	Tungro_10m_4_4K.jpg	9
	Tungro_10m_5_4K.jpg	6
<b>Total</b>		<b>69</b>

required a low-resolution (LR) image, so generating an LR image with a size of  $128 \times 128$  or downgrading four times from the HR sub-image was necessary.

After that, the training and validation process uses the ESRGAN method, using 15,000 iterations. The following is the result of the performance graph of the ESRGAN method when creating a super-resolution model.

Figure 12 shows that iterations 1000 to 15000 converged with the acquisition of PSNR values in the range 28-32 while SSIM is 0.89-0.94. Then, the loss performance of the two main ESRGAN networks also shows convergent from the beginning, so ESRGAN modelling has been considered sufficient.

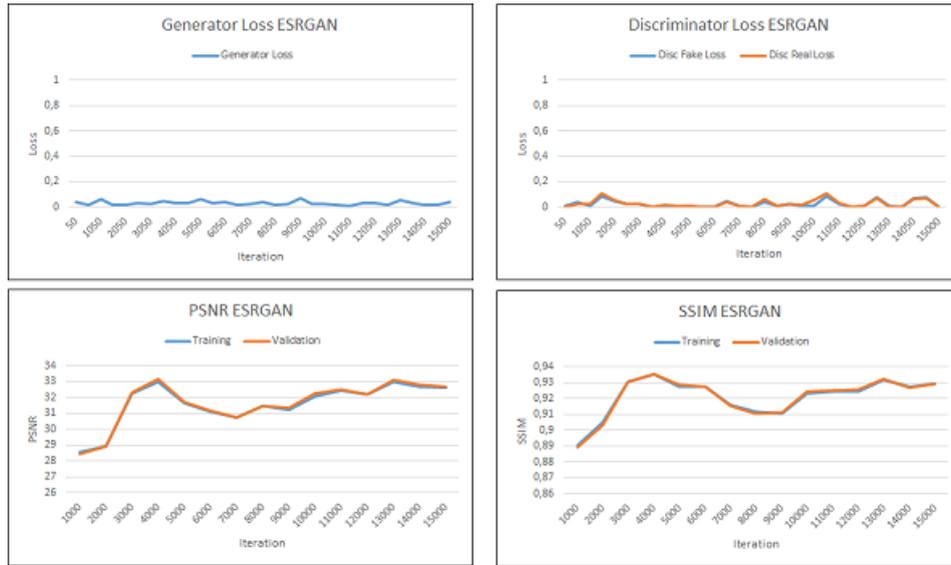


FIGURE 12. Training and Validation Performance Graph of ESRGAN Method

After obtaining the best ESRGAN model, then proceed with testing the model. When testing the model on 168 test data, the average PSNR and SSIM values were obtained: a PSNR value of 33.73 and an SSIM of 0.94. The visual differences between HR Image, LR Image, and SR Image are shown in Figure 13.



FIGURE 13. ESRGAN result images (a) HR (b) LR (c) SR

**4.3. Aerial Image Disease Classification Process Test.** There are 20 images of diseased rice, and then a sub-image extraction process is performed, where the original image is split into a size of  $512 \times 512$ .

Data splitting is done to divide the dataset into three parts, namely training data (80%) of 480 images, validation data (10%) of 60 images and test data (10%) of 60 images so that the total diseased image is 600 images. After obtaining sub-images and splitting data, the ESRGAN process is carried out on the sub-images obtained. Then, the data augmentation process is carried out in the training data so that the model obtained is increasingly trained to classify the disease. The data augmentation in this study uses translation, reflection, rotation, dilation, and brightness.

After that, the training and validation process is carried out using the CNN method using the EfficientNet-B3 architecture with 20 epochs. The results of the CNN method performance graph using the EfficientNet-B3 architecture when creating a rice disease classification model are shown in Figure 14.

After getting the best CNN model, then proceed with testing the model. The results obtained from the test of the CNN model in disease classification are shown in Table 3.

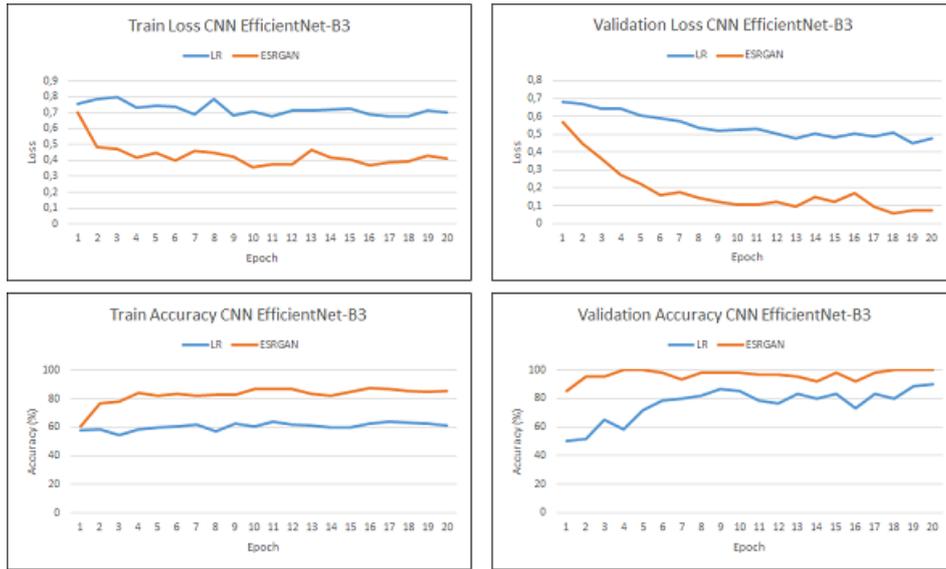


FIGURE 14. Training and Validation Performance Graph of CNN Method

TABLE 3. CNN Testing Results

CNN Data Model	Accuracy(%)	Precision(%)	Recall(%)
LR	85.00	81.82	90.00
ESRGAN	96.67	93.73	100

**4.4. Complete System Development.** The results of testing the EfficinetNet-B3 CNN architecture that has been carried out produce values on the test metrics that look quite good and the highest values are obtained from the ESRGAN data. In the previous process, namely image segmentation of the diseased area, a sub-image with a size of  $128 \times 128$  has been stored, resulting from the cropped area. Then, the sub-image is carried out in a super-resolution process, and the disease classification is performed using the best ESRGAN model obtained and the best CNN model from the previous process. The following is a visualization of the image processed for super-resolution and disease classification.

The aerial image test stage above shows the performance results of the ESRGAN model made in the previous process and the classification model produced. Figure 15 shows the visualization results for aerial images with blast disease with the file name Blast\_10m\_2.4K.jpg.



FIGURE 15. Aerial Image Visualization Results After All Processes

By applying the same process as the Blast\_10m\_2\_4K.jpg file to all aerial image files, the performance of the developed system can be evaluated. In this study, each ground-truth aerial image contains only one type of disease. Tables 4 and 5 present the final results of the disease classification system for each aerial image tested using the LR data model and the ESRGAN data model.

TABLE 4. Aerial Image Disease Classification Using The LR Data Model

Diseases	Image Name	Area	Disease Predictions	
			Blast	Tungro
Blast	Blast_10m_1_4K.jpg	8	5	3
	Blast_10m_2_4K.jpg	8	5	3
	Blast_10m_3_4K.jpg	7	4	3
	Blast_10m_4_4K.jpg	8	5	3
	Blast_10m_5_4K.jpg	5	3	2
Tungro	Tungro_10m_1_4K.jpg	4	0	4
	Tungro_10m_2_4K.jpg	9	0	9
	Tungro_10m_3_4K.jpg	5	0	5
	Tungro_10m_4_4K.jpg	9	0	9
	Tungro_10m_5_4K.jpg	6	0	6

TABLE 5. Aerial Image Disease Classification Using The ESRGAN Data Model

Diseases	Image Name	Area	Disease Predictions	
			Blast	Tungro
Blast	Blast_10m_1_4K.jpg	8	6	2
	Blast_10m_2_4K.jpg	8	5	3
	Blast_10m_3_4K.jpg	7	6	1
	Blast_10m_4_4K.jpg	8	7	1
	Blast_10m_5_4K.jpg	5	3	2
Tungro	Tungro_10m_1_4K.jpg	4	0	4
	Tungro_10m_2_4K.jpg	9	0	9
	Tungro_10m_3_4K.jpg	5	0	5
	Tungro_10m_4_4K.jpg	9	0	9
	Tungro_10m_5_4K.jpg	6	0	6

The accuracy of each data model can be calculated using the data from Tables 4 and 5. The accuracy results of each model are shown in Table 6.

TABLE 6. Result of Aerial Image Data Testing

Data Model	Accuracy(%)	Precision(%)	Recall(%)
LR	79.91	61.11	100
ESRGAN	86.96	75	100

**4.5. Analysis and Discussion.** Table 6 shows that the performance results obtained from the CNN classification method are pretty good when the super-resolution technique is performed. The accuracy value of the disease classification after using ESRGAN increased by 7.05%. These results indicate that aerial imagery's rice disease identification system requires super-resolution techniques for better results. This result is because the

ESRGAN image provides detailed characteristics of each disease, making it easier for the classification method to classify the disease.

Then the last analysis is in the calculation of the accuracy of the final results of the test aerial image Table 4 to Table 5 which shows that the two data models using any data produce a 100% classification of tungro disease because the sampling of subimage extraction of diseased areas is random so that it is possible that when blast diseases are obtained sub-images that have the characteristics of tungro.

Based on some of the above analysis, it is found that research on the identification of rice plant diseases in aerial imagery still has shortcomings, such as in the retrieval of aerial image data with blast disease, the difference is less visible from tungro disease, resulting in a less than optimal classification of blast disease. When tested on aerial imagery, the classification of tungro disease is dominant because the samples taken during the extraction of sub-images from the cropping of diseased areas may have the characteristics of tungro disease so that blast disease still needs to be maximized.

**5. Conclusion.** This research has successfully applied Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) to improve the quality of aerial images in identifying plant diseases, especially in rice plants. With several stages, namely first with the aerial image segmentation process using the Visible Atmospherically Resistant Index (VARI) method, and then with improving the quality of the aerial image using the ESRGAN method so that the aerial image has high resolution, and the final process is the identification of rice diseases using Convolutional Neural Networks (CNN) architecture *EfficientNet-B3*. Based on trials on the processes carried out in this study, the best performance resulting from the super-resolution process produces a PSNR value of 33.73 and SSIM of 0.94. The best performance for the disease classification process was obtained with an accuracy of 86.96%, precision of 75%, and a recall of 100%. These results show that the ESRGAN process successfully increased accuracy by 7.05%.

**Acknowledgment.** This research is funded by Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat, Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi based on Kontrak Induk Pelaksanaan Program Bantuan Operasional Perguruan Tinggi Negeri Penelitian Fundamental. Nomor Kontrak Induk: 038/E5/PG.02.00.PL/2024, date 11 June 2024. Nomor Kontrak Peneliti: 1750/PKS/ITS/2024, date 12 June 2024. We would like to thank the Indonesian farmers in Tulungagung City.

## REFERENCES

- [1] R. K. Dubey, and D. K. Choubey, Efficient prediction of blast disease in paddy plant using optimized support vector machine, *IETE Journal of Research*, vol.70, no.4, pp.3679–3689, 2023.
- [2] G. L. Hanopol, and J. C. D. Cruz, Rice tungro disease detection based on light absorption assessment using machine learning, *In 2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC)*, pp. 145-150, 2022.
- [3] X. Zhang, H. Dong, L. Gong, X. Cheng, Z. Ge, and L. Guo, Multiple paddy disease recognition methods based on deformable transformer attention mechanism in complex scenarios, *International Journal of Computers and Applications*, vol.45, no.10, pp.660-672, 2023.
- [4] K. S. Iskandar, B. N. Afiza, and B. N. Atiqah, Introduction to drone technology for agriculture purposes: a brief review, *International Journal of Agriculture and Biological Sciences*, vol.4, no.4, pp.13–23, 2020.
- [5] A. A. Haykir, and I. Oksuz, Super-resolution with generative adversarial networks for improved object detection in aerial images, *Information Discovery and Delivery*, vol.51, no.4, pp.349–357, 2022.

- [6] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, Esrgan: Enhanced super-resolution generative adversarial networks, *In Proceedings of the European conference on computer vision (ECCV) workshops*, 2018.
- [7] L. Zha, Y. Shi, and J. Wen, A Lightweight Image Super-Resolution Network Based on ESRGAN for Rapid Tomato Leaf Disease Classification, *In The International Conference on Image, Vision and Intelligent Systems (ICIVIS 2021)*, Singapore: Springer Nature Singapore, pp.97–110, 2022.
- [8] B. Setiyono, D. R. Sulistyaningrum, A. F. Pratama, and R. N. R. Wijaya, A Modification of the Temporal Group Attention Method on Super-Resolution Video for Vehicle Number Plate Detection, *J Inf Hiding Multim Signal Process*, vol.15, no.3, pp.156–165,2024.
- [9] M. F. Rabbi, M. N. Sultan, M. Hasan, and M. Z. Islam, Tribal Dress Identification using Convolutional Neural Network, *J Inf Hiding Multim Signal Process*, vol.14, no.3, pp.72–80,2023.
- [10] A. U. Rehman, and C. F. Lee, CNN-Based Intelligent Disease Detection and Identification Technique through Chest X-rays, *J Inf Hiding Multim Signal Process*, vol.15, no.4, pp.319–332,2024.
- [11] W. Albattah, A. Javed, M. Nawaz, M. Masood, and S. Albahli, Artificial intelligence-based drone system for multiclass plant disease detection using an improved efficient convolutional neural network, *Frontiers in Plant Science*, vol.13, 2022.
- [12] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour, and N. A. Alajlan, Classification of remote sensing images using efficientnet-b3 cnn model with attention, *IEEE access*, vol.9 , pp.14078—14094, 2021.
- [13] K. Yamamoto, T. Togami, and N. Yamaguchi, Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture, *Sensors*, vol.17, no.11, 2557, 2017.
- [14] A. Moghimi, C. Yang, and J. A. Anderson, Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat, *Computers and Electronics in Agriculture*, vol.172, 105299, 2020.
- [15] M. G. Selvaraj, A. Vergara, F. Montenegro, H. A. Ruiz, N. Safari, D. Raymaekers, W. Ocimati, J. Ntamwira, L. Tits, A. B. Omondi, and G. Blomme, Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.169, pp.110–124, 2020.
- [16] S. Chayhard, V. Mantachitra, K. Nualchawee, and A. Buranapratheprat, Application of aerial photography with visible atmospherically resistant index by using unmanned aerial vehicles for seagrass bed classification in Kung Krabaen Bay, Thailand, *PeerJ Preprints*, vol.6, e27407v1, 2018.
- [17] V. Nascimento, R. Laroca, J. D. A. Lambert, W. R. Schwartz, and D. Menotti, Super-resolution of license plate images using attention modules and sub-pixel convolution layers, *Computers and Graphics*, vol.113, pp.69–76, 2023.
- [18] Z. Wu, and P. Ma, ESRGAN-based DEM super-resolution for enhanced slope deformation monitoring in lantau island of Hong Kong', *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol.43, pp.351–356, 2020.
- [19] C. Ma, C. Y. Yang, X. Yang, and M. H. Yang, Learning a no-reference quality metric for single-image super-resolution, *Computer Vision and Image Understanding*, vol.158, pp.1–16, 2017.
- [20] M. Tan, and Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks', *In International conference on machine learning*, PMLR, pp.6105–6114, 2020.
- [21] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, Fundamental concepts of convolutional neural network, *Recent trends and advances in artificial intelligence and Internet of Things*, pp.519–567, 2020.
- [22] D. Krstinić, M. Braović, L. Šerić, and D. Božić-Štulić, Multi-label classifier performance evaluation with confusion matrix, *Computer Science and Information Technology*, vol.10, no.8, pp.1-14, 2020.