SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Student Research Exploration

https://shmpublisher.com/index.php/josre
p-ISSN xxxx-xxxx | e-ISSN xxxx-xxxx

Detection and prediction of rice plant diseases using convolutional neural network (CNN) method

Reyhan Dzaki Sheva[1], Halimah Salsabila[2], Kusuma Ratna Pratiwi[3]
[1, 2]Department of Computer Science, Universitas Negeri Semarang, Indonesia
[3]Faculty of Animal and Agricultural Sciences, Universitas Dipoengoro

Article Info

Article history:

Received November 27, 2023
Revised January 15, 2024
Accepted January 20, 2024

Keywords:

Detection
Rice plant disease
Machine learning
Image processing
CNN

ABSTRACT

Rice is a basic staple food in many Asian countries and is generally irreplaceable. Rice accounts for almost half of Asia food expenditure. Rice is too a crop that is prone to plant disease. It can appear and cause a decline in the quality of rice. However, constant monitoring of the rice fields can prevent the infection of the disease. Therefore, detection and prediction of rice plant diseases is one of the topics that will be discussed in this research. The purpose of this research is to help farmers to quickly pinpoint the disease of rice plants and take care of it properly. The methods used in this paper is researching and redesigning the previous attempt to hopefully make it better and more accurate. We will be using Convolutional Neural Network (CNN) models VGG16 as our algorithm. The results are that our proposed method has more accuracy than previous research using a similar dataset. The novelty of this paper is the increased accuracy of rice plant disease detection.

1. Introduction

Rice, wheat, and maize are three of the main crops that are widely consumed around the world. The more popular choice of main crops to consume in low- and lower-middle-income countries is rice. In many Asian countries, rice is a basic staple food and is generally irreplaceable [1]. For people of all financial backgrounds in Asia, rice was considered an important basic food. It is the most

[1] ***Corresponding Author:***
Reyhan Dzaki Sheva Pahlawanto,
Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang, Indonesia,
Sekaran, Gunungpati, Kota Semarang, Indonesia.
Email: reyhandzakisp@students.unnes.ac.id
DOI: https://doi.org/10.52465/josre.v2i1.254

important energy source and essential to the food security of many people around the world.

Plant diseases are one of the reasons for declining quantity and quality of agricultural produce. A decline in either side could have a direct impact on the country's overall crop production and could cause severe damage to the food supply [2]. More often than not, diseases can occur in any crop and at any moment. One crop that is prone to disease is rice. As explained earlier, rice is one of the most consumed foodstuffs around the world, especially on the Asian continent. If rice plants are not properly cared for, various diseases will appear on rice plants that can cause a decline in quality of crops and rice [3]. However, constant monitoring could prevent that problem. Therefore, detection and prediction of rice plant diseases is one of the topics that will be discussed in this research.

There are several similar research that has been conducted on image processing. The first example is research that has been conducted by Harshadkumar B. P., Jitesh P. S., and Vipul K. D. [4] about rice plant disease detection and classification using Support Vector Machine (SVM). They created their own dataset using images from a rice field in a village called Sherta near Gandhinagar, Gujarat, India. and experimented with a few image processing techniques until they came up with the most effective techniques at the time of the research. Another research about this similar topic is done by Rismiyati and Ardytha Luthfiarta [5] about salak fruit quality classification using VGG16 transfer learning architecture that is able differentiate between a bad quality salak and a good quality salak. Another study that was related to the topic is about automatic classification and detection of leaf diseases using software solutions that have enabled faster and more accurate solutions and it was conducted by H. Al-Hiary with other writers [6]. Another research has been conducted by Santanu Phadikar and others about using feature selection and rule generation techniques as a rice diseases classification utilizing Rough Set Theory (RST) [7]. Another study that was related to the topic is about using AlexNet as a rice leaf disease detection technique and it is written by Md. Mafiul Hasan Matin, Amina Khatun, with other writers [8]. Another research about this similar topic is done by Halil Durmuú, Ece Olcay Güneú, Mürvet Korco and it′s about tomato leaf disease detection by using deep learning techniques such as AlexNet and SqueezeNet [9]. Another study that was related to the topic is about using deep neural networks to automatically detect UAV images of tobacco plants and it was conducted by Zhun Fan with other writers [10]. Another similar research was conducted by Zhenzhen Song with other writers and it′s about the VGG16 method to detect kiwifruit in field images [11]. Another research has been conducted by Zhong Qu with other writers and it talks about using the VGG16 network model to detect cracks in concrete pavement [12]. Another study that has been conducted on this topic is about batik image classification using VGG16 based on random forest, and it was conducted by D. M. S. Arsa and A. A. N. H. Susila [13]. The main research that will be used as a reference will be discussed below.

The rice plant disease recognition system can be built using the Support Vector Machine (SVM), Local Binary Pattern (LBP), Convolutional Neural Network (CNN), and other algorithms. This research has previously been conducted by Harshadkumar and others using the Support Vector Machine (SVM) algorithm. By using their own dataset, they get an accuracy of 93.33% on the training dataset and an accuracy of 73.33% on the test dataset. The dataset they use is available on Kaggle, therefore the author will be using the same dataset with a different algorithm method.

This research attempts to test the Machine Learning and Image Processing concepts to figure out the solution of rice disease detection and prediction, more specifically using Convolutional Neural Network (CNN) using the same dataset that the main reference paper used. Convolutional Neural Network (CNN) is the most widely used feature extraction network in various studies for the task of detecting and classifying plant diseases [14]. For the rice disease dataset, the most recurrent diseases that infected rice plants are bacterial leaf blight, brown spot, and leaf smut. Each disease has format characteristics such as different shapes, sizes, and colours. Sometimes farmers become confused and cannot make the right decision to deal with the disease. Photographing infected leaves and collecting data and details about the disease is one of the solutions to avoid yield loss due to the infection. Therefore, image processing can be used to help detect diseases more accurately and in less time [15]. The essence of this research is to provide a solution for automatic recognition of rice diseases using these images. The author tries to address the problem in this research.

2. Method

The design of the research method that will be carried out in this research is shown in Figure 1.

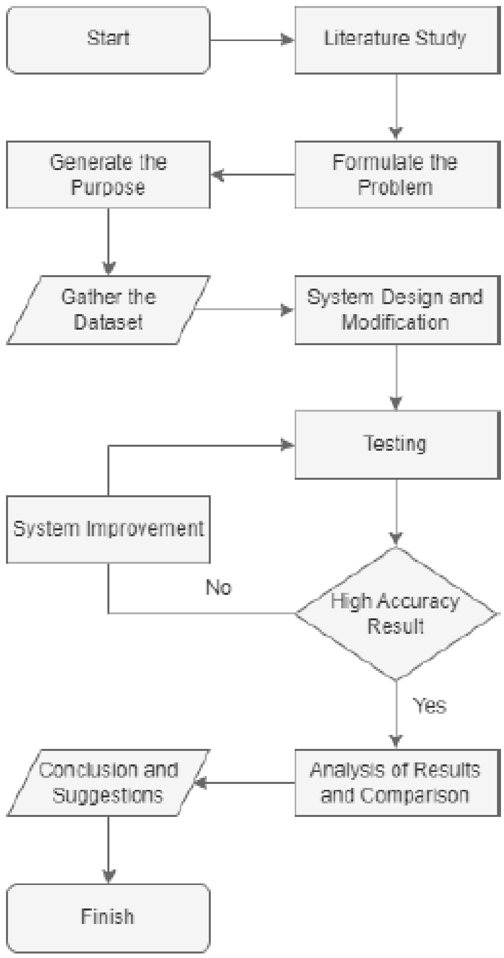

Figure 1. Diagram of the research method

First, the researcher will conduct a literature study. This is necessary so that researchers understand the problem to be studied and what needs to be done to achieve the desired solution in this research. Then the researcher will formulate a problem, where the problem that will be discussed by the researcher is about detecting and predicting diseases in rice plants. This is a particularly relevant issue as Indonesia is one of the world's largest rice producers [16]. Indonesia is also the world's largest rice consumer. Due to the unpredictable weather, many rice plants are susceptible to disease. Another problem is that determining disease on rice leaves by human senses often takes time because the quality is often unstable [17]. Therefore, the care and maintenance of rice plants is something that is very well considered, so that rice plants can produce decent and quality rice. After that, the researcher will generate a purpose for this paper. The purpose of this research is to hopefully produce a precise and accurate rice plant disease detection and prediction system, so that farmers can be helped with their work and could take early action to protect their crops [18].

The next step is to gather a dataset to use in this research. The dataset that will be used in this research is the Rice Leaf Diseases Dataset obtained from Kaggle [19]. This dataset contains a total of 120 images of rice leaves infected with diseases in

JPG extension format. The images are classified into three classes according to the type of disease, namely bacterial leaf blight, brown spot, and leaf smut. Images of the three types of disease can be seen in Figure 2. There are forty images in each class. The images have various sizes with all of them being RGB types of images. Some images have their background removed while others are kept intact. The author will separate the folder into a training folder and a test folder. Details of the dataset obtained can be seen in table 1.

Table 1. Division of the rice leaf diseases dataset

No.	Type of data	Division	
		Total	Each Classes
1	All Dataset	120 Images	40 Images
2	Training Dataset	105 Images	35 Images
3	Test Dataset	15 Images	5 Images

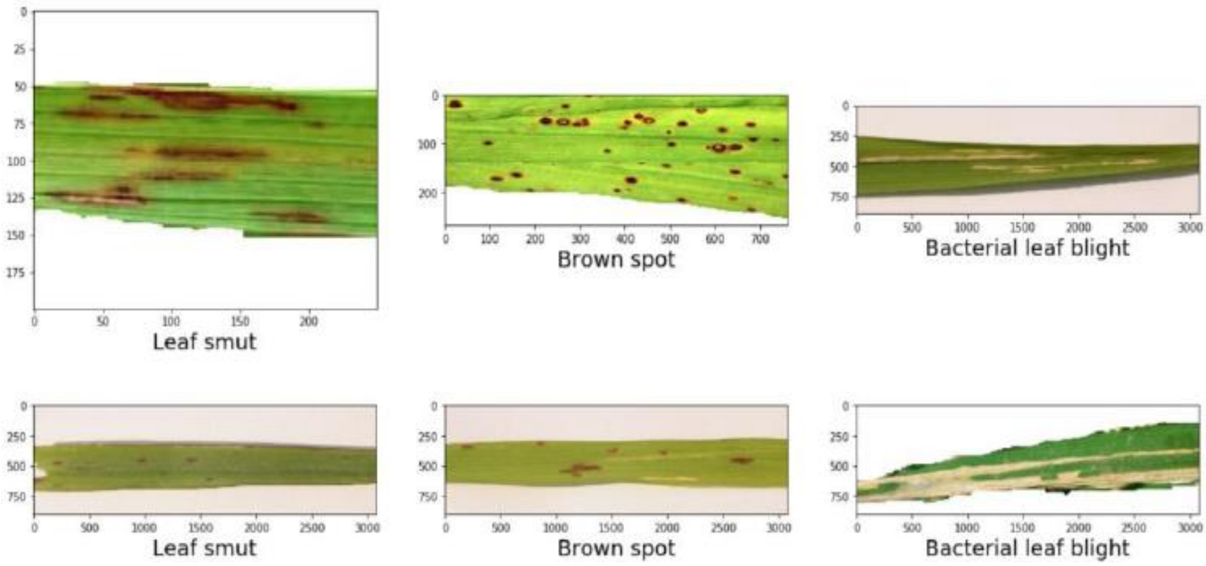

Figure 2. Preview of the content of the rice leaf diseases dataset

The next step is designing and modifying the system. The following is a system design that will be built in the form of a flowchart in Figure 3.

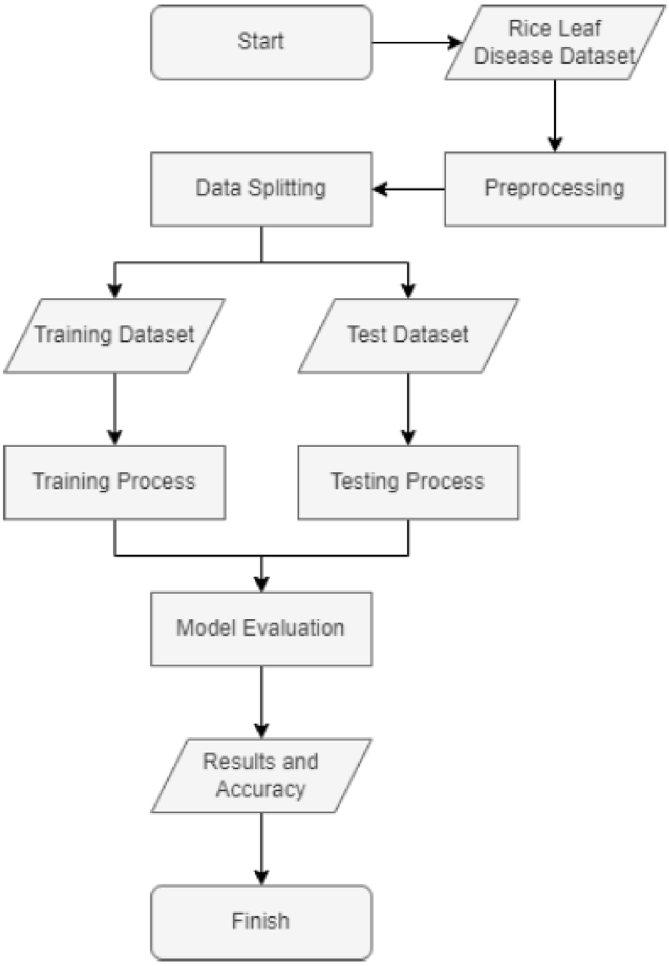

Figure 3. Flowchart of the system design

Preprocessing is done by resizing the image. The image size is changed to 224x224x3 pixels. The image is not converted to grayscale image because color features are important features in this recognition. After that, the image is split up into 87.5% training data and 12.5% testing data. The rice plant disease dataset is used as input. The data was then subjected to a preprocessing process. After that, the data was divided into 87.5% training data and 12.5% testing data. The training data was then trained using VGG16 with the Tensorflow library and Keras API. TensorFlow is an open-source machine learning software library developed by Google that enables you to deploy computations on CPUs or GPUs using graphs flow data [20], while Keras API is a high-level neural network API originally written and available in Python that works on TensorFlow [21]. In this process, there are several parameters used, namely epoch and batch size. Epoch determines the number of iterations that must be performed on the data set, while batch size determines the amount of data that is trained in one iteration. The final model results are then evaluated using testing. The evaluation results are in the form of training accuracy value and validation accuracy of the evaluated model observed for each research scenario.

For this kind of task, various CNN-based models have been developed, such as AlexNet, VGGNet, GoogleNet, and others [22]. The algorithm that will be used in this research is Convolutional Neural Network (CNN) with the VGG16 architecture model. VGG16 is a CNN model that uses convolutional layers with small (3×3) convolutional filter specifications. It aims to significantly increase the depth of existing CNN architectures, which typically have 16 or 19 layers [23]. VGG16 was proposed by K. Simonyan and his friend at the Oxford University in their work on deep convolutional networks for large-scale image recognition [24]. This model consists of 16 hidden layers which contain thirteen convolutional layers, two fully connected layers and one classifier layer. The kernel size of all convolutional layers is 3x3. The main difference between each convolutional layer is the number of filters in each layer. The first two convolutional layers have a total of 64 filters. Layers 3 and 4 have a total of 128 kernels. Similarly, the other convolutional layers have different numbers of filters which are 256 and 512. The maximum 2x2 Max Pooling occurs after a particular convolutional layer. The output of the last assembly is combined with the fully connected layers and finally with the classifier to determine the image class. For a clearer Architecture of VGG16, see Figure 4.

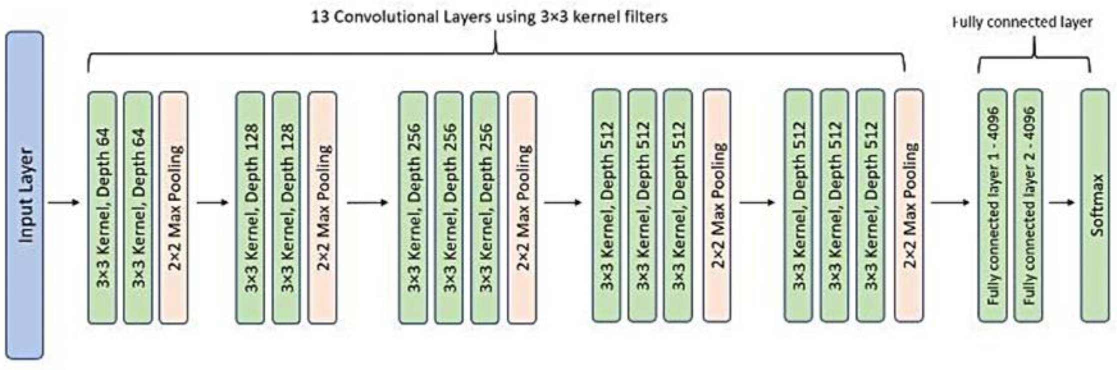

Figure 4. Architecture of VGG16 [25]

3. Results and Discussion

The system model was executed through Google Colab, which took approximately 5-10 minutes. The dataset used is Rice Leaf Disease Dataset which consists of 120 images divided into 3 classes, namely Bacterial leaf blight, Brown spot, and Leaf smut. Which can be seen in Figures a, b, and c.

(a)

(b)

(c)

Figure 5. An example of bacterial leaf blight (a), brown spot (b), and leaf smut (c)

The tests were conducted by utilizing several parameters, namely batch size of 32 and epoch of 100. The tested model is Convolutional Neural Network (CNN) with VGG16 model, then compared with Support Vector Machine (SVM) architecture model from research conducted by Harshadkumar B. P. and others. The result obtained when testing the built system is the accuracy of the model used. The accuracy results of each model can be seen in table 2.

Table 2. Training and test accuracy and loss of each model

No.	Model	Training		Test	
		Accuracy	Loss	Accuracy	Loss
1	SVM Model	0.9333	-	0.7333	-
2	Proposed VGG16 Model	~1.0000	0.0456	0.9333	0.1040

The results obtained from one of the fitting models performed on the training dataset obtained accuracy of approximately ~1.0000 and on the test dataset obtained accuracy of 0.9333.

The author argues that the reason for obtaining these results is the presence of Image Augmentation. This is a technique that could be used to increase the size of the training dataset artificially. It creates images that were modified within the dataset. With this, CNN models could be trained on more data that can result in more skilled models, and this technique could generate a variety of images that can enhance the model's ability to transfer what it has learned to new images. Figure 6 displays the accuracy graph and figure 7 displays the loss of the training dataset and testing dataset.

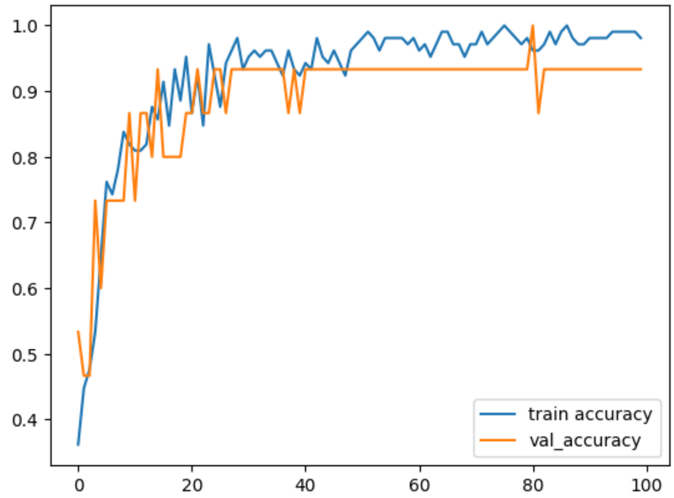

Figure 6. Graph of training and testing accuracy

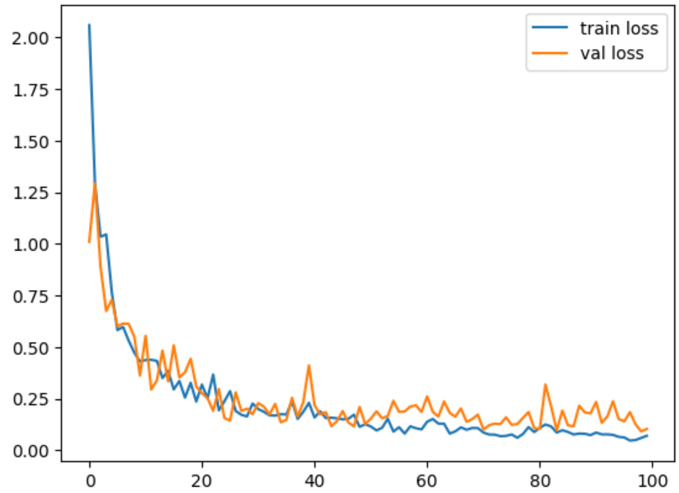

Figure 7. Graph of training and testing loss

For comparison, the author uses a journal written by Harshadkumar B. P., Jitesh P. Shah. and Vipul K. D. entitled "Detection and Classification of Rice Plant Diseases". In the journal, researchers used K-Means Clustering and Support Vector Machine (SVM). Table 3 shows a comparison of the results of previous research with this research.

Table 3. Training and test accuracy of each model

No.	Research	Dataset	Model	Accuracy	
				Training	Test
1	Harshadkumar B. P., Jitesh P. S., and Vipul K. D.	Rice Leaf Diseases Dataset	SVM Model	93.33%	73.33%
2	This research		Proposed VGG16 Model	~100%	93.33%

The results obtained by the previous journal are 93.33% accuracy in the training dataset and 73.33% accuracy in the test dataset. While in this research, the author managed to get an accuracy of ~100% in the training dataset and 93.33% in the test dataset, although the accuracy in the training dataset is not completely consistent, the author is quite satisfied with this result. In addition, the predictions produced are quite satisfactory because the results obtained are in accordance with the available datasets.

4. Conclusion

Based on the results of the research conducted, CNN and VGG16 algorithms provide more accurate results of disease detection and prediction in rice plants compared to K-Means Clustering and SVM algorithms. This can be seen from the greater accuracy value in the CNN and VGG16 algorithms. In previous research using the K-Means Clustering and SVM algorithms, 93.33% accuracy was obtained in the training dataset and 73.33% accuracy in the test dataset. While in this research, the author managed to get an accuracy of ~100% in the training dataset and 93.33% in the test dataset. This shows that CNN and VGG16 algorithms are able to provide more accurate results compared to previous algorithms. Thus, CNN models VGG16 algorithms can be used as an option in detecting and predicting diseases in rice plants, so as to facilitate the work of farmers in caring for rice plants. The evaluation that the author can provide is to use a dataset that is larger in size and contains more images, to better understand how good the performance of this algorithm is. In addition, the algorithm may be researched more deeply so that it can produce consistent and high accuracy in a short time too, because there are some differences in accuracy from some of the test results conducted.

REFERENCES

[1] Z. Deng, H. Hung, M. T. Carson, A. A. Oktaviana, B. Hakim, and T. Simanjuntak, "Validating earliest rice farming in the Indonesian Archipelago," *Sci. Rep.*, vol. 10, no. 1, p. 10984, Jul. 2020, doi: 10.1038/s41598-020-67747-3.

[2] E. Acar, O. F. Ertugrul, E. Aldemir, and A. Oztekin, "Automatic identification of cassava leaf diseases utilizing morphological hidden patterns and multi-feature textures with a distributed structure-based classification approach," *J. Plant Dis. Prot.*, vol. 129, no. 3, pp. 605–621, Jun. 2022, doi: 10.1007/s41348-022-00583-x.

[3] G. S. Laha *et al.*, "Importance and Management of Rice Diseases: A Global Perspective," in *Rice Production Worldwide*, Cham: Springer International Publishing, 2017, pp. 303–360. doi: 10.1007/978-3-319-47516-5_13.

[4] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, "Detection and classification of rice plant diseases," *Intell. Decis. Technol.*, vol. 11, no. 3, pp. 357–373, Aug. 2017, doi: 10.3233/IDT-170301.

[5] R. Rismiyati and A. Luthfiarta, "VGG16 Transfer Learning Architecture for Salak Fruit Quality Classification," *Telematika*, vol. 18, no. 1, p. 37, Mar. 2021, doi: 10.31315/telematika.v18i1.4025.

[6] H. Al-Hiary, B. S. Ahmad, M. Reyalat, M. Braik, and Z. ALRahamneh, "Fast and Accurate

Detection and Classification of Plant Diseases," *Int. J. Comput. Appl.*, vol. 17, no. 1, pp. 31–38, 2011, doi: 10.5120/2183-2754.

[7] S. Phadikar, J. Sil, and A. K. Das, "Rice diseases classification using feature selection and rule generation techniques," *Comput. Electron. Agric.*, vol. 90, pp. 76–85, Jan. 2013, doi: 10.1016/j.compag.2012.11.001.

[8] M. M. H. Matin, A. Khatun, M. G. Moazzam, and M. S. Uddin, "An Efficient Disease Detection Technique of Rice Leaf Using AlexNet," *J. Comput. Commun.*, vol. 08, no. 12, pp. 49–57, 2020, doi: 10.4236/jcc.2020.812005.

[9] H. Durmus, E. O. Gunes, and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," in *2017 6th International Conference on Agro-Geoinformatics*, Aug. 2017, pp. 1–5. doi: 10.1109/Agro-Geoinformatics.2017.8047016.

[10] Z. Fan, J. Lu, M. Gong, H. Xie, and E. D. Goodman, "Automatic Tobacco Plant Detection in UAV Images via Deep Neural Networks," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 3, pp. 876–887, Mar. 2018, doi: 10.1109/JSTARS.2018.2793849.

[11] Z. Song, L. Fu, J. Wu, Z. Liu, R. Li, and Y. Cui, "Kiwifruit detection in field images using Faster R-CNN with VGG16," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 76–81, 2019, doi: 10.1016/j.ifacol.2019.12.500.

[12] D. M. S. Arsa and A. A. N. H. Susila, "VGG16 in Batik Classification based on Random Forest," in *2019 International Conference on Information Management and Technology (ICIMTech)*, Aug. 2019, pp. 295–299. doi: 10.1109/ICIMTech.2019.8843844.

[13] Z. Qu, J. Mei, L. Liu, and D.-Y. Zhou, "Crack Detection of Concrete Pavement With Cross-Entropy Loss Function and Improved VGG16 Network Model," *IEEE Access*, vol. 8, pp. 54564–54573, 2020, doi: 10.1109/ACCESS.2020.2981561.

[14] S. Mishra, R. Sachan, and D. Rajpal, "Deep Convolutional Neural Network based Detection System for Real-time Corn Plant Disease Recognition," *Procedia Comput. Sci.*, vol. 167, pp. 2003–2010, 2020, doi: 10.1016/j.procs.2020.03.236.

[15] T. Islam, M. Sah, S. Baral, and R. Roy Choudhury, "A Faster Technique on Rice Disease Detectionusing Image Processing of Affected Area in Agro-Field," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Apr. 2018, pp. 62–66. doi: 10.1109/ICICCT.2018.8473322.

[16] A. Bashir and S. Yuliana, "Identifying factors influencing rice production and consumption in Indonesia," *J. Ekon. Pembang. Kaji. Masal. Ekon. dan Pembang.*, vol. 19, no. 2, Jan. 2019, doi: 10.23917/jep.v19i2.5939.

[17] J. Pardede, B. Sitohang, S. Akbar, and M. L. Khodra, "Implementation of Transfer Learning Using VGG16 on Fruit Ripeness Detection," *Int. J. Intell. Syst. Appl.*, vol. 13, no. 2, pp. 52–61, Apr. 2021, doi: 10.5815/ijisa.2021.02.04.

[18] T. Daniya and S. Vigneshwari, "A review on machine learning techniques for rice plant disease detection in agricultural research," *Int. J. Adv. Sci. Technol.*, vol. 28, no. 13, pp. 49–62, 2019.

[19] VBOOKSHELF, "Rice Leaf Diseases Dataset," *Kaggle*, 2019. https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases

[20] S. A. Sanchez, H. J. Romero, and A. D. Morales, "A review: Comparison of performance metrics of pretrained models for object detection using the TensorFlow framework," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 844, p. 012024, Jun. 2020, doi: 10.1088/1757-899X/844/1/012024.

[21] N. K. Manaswi, "Understanding and Working with Keras," in *Deep Learning with Applications Using Python*, Berkeley, CA: Apress, 2018, pp. 31–43. doi: 10.1007/978-1-4842-3516-4_2.

[22] J. Lu, L. Tan, and H. Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," *Agriculture*, vol. 11, no. 8, p. 707, Jul. 2021, doi: 10.3390/agriculture11080707.

[23] H.-C. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image

recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015, doi: https://doi.org/10.48550/arXiv.1409.1556.
[25] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *Int. J. Sci. Res. Publ.*, vol. 9, no. 10, p. p9420, Oct. 2019, doi: 10.29322/IJSRP.9.10.2019.p9420.