INTENSIF, Vol.8 No.1 February 2024
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

Augmented Rice Plant Disease Detection with Convolutional Neural Networks

Received:
19 September 2023
Accepted:
18 January 2024
Published:
1 February 2024

[1*]Hairani, [2]Triyanna Widiyaningtyas
[1-2]Electrical Engineering and Informatics, Universitas Negeri Malang
E-mail: [1]hairani.2305349@students.um.ac.id, [2]triyannaw.ft@um.ac.id
*Corresponding Author

Abstract—The recognition and classification of rice plant diseases require an accurate system to generate classification data. Types of rice diseases can be identified in several ways, one of which is leaf characterization. One method that has high accuracy in identifying plant disease types is Convolutional Neural Networks (CNN). However, the rice disease data used has unbalanced data which affects the performance of the method. Therefore, the purpose of this research was to apply data augmentation to handle unbalanced rice disease data to improve the performance of the Convolutional Neural Network (CNN) method for rice disease type detection based on leaf images. The method used in this research is the CNN method for detecting rice disease types based on leaf images. The result of this research was the CNN method with 100 epochs able to produce an accuracy of 99.7% in detecting rice diseases based on leaf images with a division of 80% training data (2438 data) and 20% testing data (608 data). The conclusion is that the CNN method with the augmentation process can be used in rice disease detection because it has very high accuracy.
Keywords— Augmentation Data; Deep Learning; Convolutional Neural Network; Rice Plant Disease Detection

Corresponding Author:
Hairani,
Electrical Engineering and Informatics,
Universitas Negeri Malang,
Email: hairani.2305349@students.um.ac.id
Orcid ID: http://orcid.org/0000-0002-6756-5896

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

I. INTRODUCTION

Indonesia is known as the world's third-largest producer and consumer of rice [1]. Data from the Statistics Indonesia shows that around 28 million Indonesians in 2016 were farmers, some living below the poverty line [2]. Activities that increase rice productivity will impact millions of rice farmers in Indonesia [3]. Farmers are estimated to lose 10 – 30% of their rice yearly due to pests and diseases [4]. Knowledge about pests and rice plant diseases significantly increases farmers' income [5]. Therefore, it is necessary to develop a system to recognize and classify rice plant diseases to help Indonesian rice farmers [6]. The recognition and classification of rice plant diseases require an accurate system to generate classification data [7]. Types of rice diseases can be identified in several ways, one of which is leaf characteristics [8].

Several previous studies have conducted rice disease detection using various approaches, such as research [9] using the Random Forest method for medicinal plant disease detection with an accuracy of 98.97%, precision 99.42%, recall 98.89%, and F-measure 99.15%. Research [10] applies genetic feature selection with the SVM method for variety detection to get an accuracy of 87.67% and an AUC of 93%. Research [11] uses the Deep Convolutional Neural Network (DCNN) method to detect rice disease types with precision 0.962, recall 0, 0.9617, specificity 0.9921, and F1-score 0.9616. Research [12] uses the CNN method to detect rice disease types as much as 4955 data with an accuracy of 96%. Research [13] utilizes deep learning technology for apple fruit type detection of as many as 1990 images with an accuracy of 98%. Research [14] employs the AlexNet transfer learning method with data as much as 600 images to get an accuracy of 91.23%. Research [15] uses feature extraction techniques for detecting rice plant diseases, namely bacterial leaf blight, brown spot, and leaf, with an accuracy of 86%. Research [16] uses DenseNet121, DenseNet169, and DenseNet201 methods to detect rice plant diseases such as bacterial leaf blight, brown spot, and leaf. Accuracy using DenseNet121 was 91.67%, DenseNet169 was 90%, and DenseNet201 was 88.33%. The three models require 24 seconds of computing time. Research [17] uses the CNN method to classify Oryza sativa rice types with an accuracy of 95%. Research [18] used the Gray Level Co-occurrence Matrix (GLCM) method for feature extraction in diagnosing rice plant disease types with an accuracy of 90%.

Research [19] uses Fuzzy C-Means and Genetics methods for clustering types of rice plant diseases with a precision of 65%. Research [20] utilizes the SVN method for detecting rice plant diseases based on leaf images with 91.3% accuracy, 90.72% sensitivity, 91.88 specificity, and 92% precision. Research [21] employs the Convolutional Neural Network method for detecting rice plant diseases based on leaf images with an accuracy of 95%. Research [22] uses the VGG16 method for disease detection in rice plants with an accuracy of 78%. Research [23] uses deep

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

learning methods with data of as much as 33,026 images with 91% accuracy. Research [24] uses the XGBOOST method for disease detection in rice plants based on images with an accuracy of 93%. Research [25] uses the Convolutional Neural Network (CNN) method for disease detection in rice with as many as 3355 images with an accuracy of 93%. However, previous research has not resolved the problem of unbalanced data on rice diseases which affects the performance of the method.

Many studies are now being conducted on using Deep Learning technology to diagnose rice diseases [26] [27]. This research uses deep learning methods that can be used as a rice-plant disease detection technique. This research to implement the deep learning method Convolutional Neural Network for detecting rice plant disease types. Research on disease detection in other plants, such as tomatoes and chili, has been done, but research on disease detection in rice plants is still small [28].

Some gaps can be addressed based on problems that have not been resolved by previous research, such as (1) the accuracy is still not optimal and can be improved, (2) the dataset of rice plant disease types used is unbalanced data, which affects the accuracy of the method. Therefore, the purpose of this research was to apply data augmentation to handle unbalanced rice disease data to improve the accuracy of the Convolutional Neural Network (CNN) method for rice disease type detection based on leaf images. Unbalanced image data is balanced using data augmentation, and the CNN method is used for the classification of rice disease types based on data that has been balanced. CNN is used to identify types of plant diseases because it has high accuracy.

II. RESEARCH METHOD

This research consists of several stages, as seen in Figure 1.

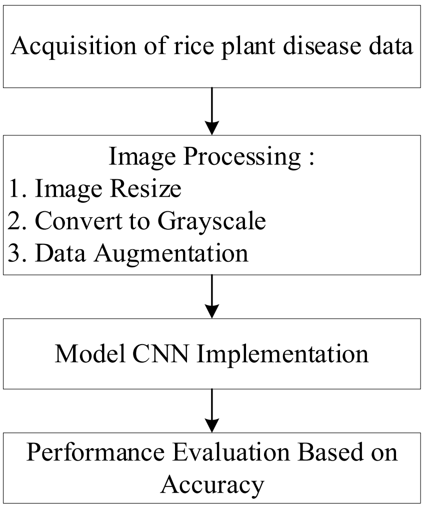


Fig 1. Research Flow

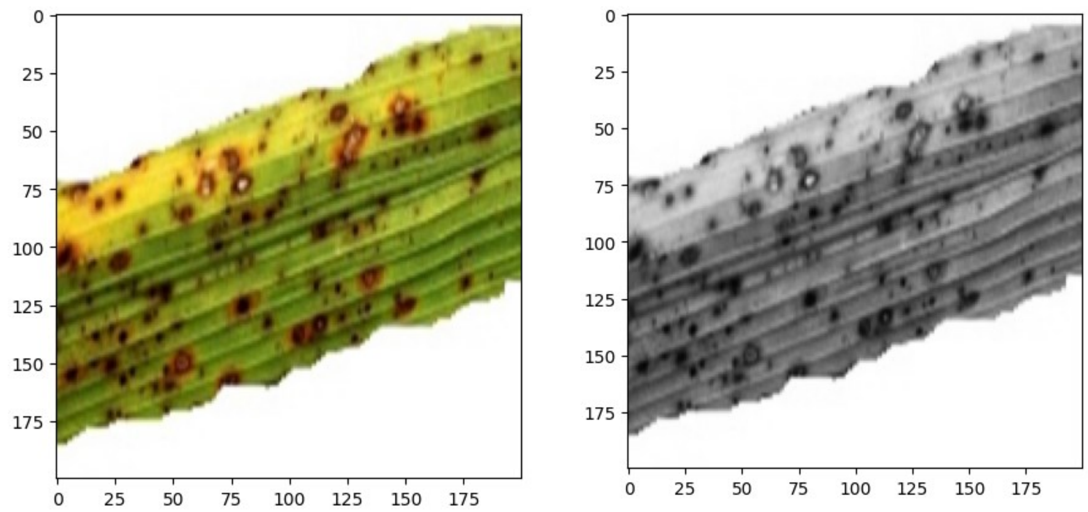
INTENSIF, Vol.8 No.1 February 2024
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

A. Acquisition of Rice Plant Disease Data

Collecting data on rice plants based on leaf images obtained from Kaggle. Rice leaf image data is obtained from Kagle with a label of 3 diseases: Bacterial leaf blight, Brown spot, and Leaf smut. The data is 120 images, and each disease class consists of 40 images.

B. Image Processing

Image processing proccess, which is used to improve the quality of rice plant leaf images that can reduce overfitting and optimize the performance of the CNN model. The techniques used in the image processing stage are resizing, converting RGB to grayscale images, and data augmentation. Resize image serves to homogenize the size of an image to reduce computation time [29]. This research uses an image size of 200 * 200. Changing an RGB image to a grayscale image makes the execution time and complexity lower, as the complexity level can affect the model's performance in image recognition [30]. Data augmentation is needed to add more data so that it is more varied so that the CNN method is more optimal in recognizing rice disease data [31]. The data augmentation process uses several techniques so that the augmented data is not 100% similar to the original data. The augmentation techniques used are Rescale, Shear, Zoom, Rotation, and Flip Horizontal like figure 2.



(a) (b)

Fig 2. Result Image Grayscale; (a). Image Before Grayscale, (b). Image After Grayscale

C. Model CNN Implementation

Implementing the CNN method for leaf image-based rice plant disease detection using 80% training data and 20% testing data. CNN methods generally have convolutional layers, pooling layers, and fully connected layer processes in image prediction or classification processes. The convolution process extracts features and learns the representation of informative features from the input image. The features generated from the convolution process are pooled. The pooling layer process reduces the feature space size using the Max Pooling or Average Pooling method

INTENSIF, Vol.8 No.1 February 2024
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

to reduce computation time. The filter sizes used in the pooling process are usually 2*2 and 5*5. Next is the Flatten process. The Flatten process changes the size of the m*n matrix to 1 dimension in the form of input vectors. A more precise explanation of the work process of the CNN method can be seen in Figure 3.

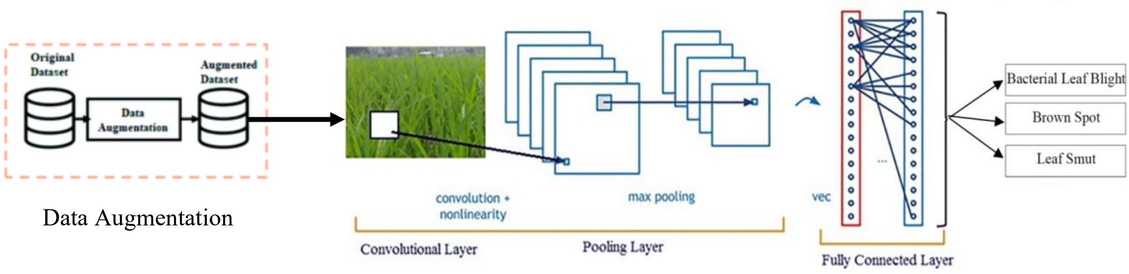


Fig 3. Work Process of the CNN Method

D. Performance Evaluation

Evaluating the performance of the CNN method for leaf image-based rice plant disease detection based on accuracy. Accuracy is used to evaluate the ability of the CNN model to identify rice disease classes from the entire rice dataset. The formula used to calculate accuracy is as in Equation (1) [32], [33] .

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

III. RESULT AND DISCUSSION

Data on rice plant diseases was obtained from the Kaggle site with a total of 120 data and had a label of 3 diseases, namely Bacterial leaf blight, Brown spot, and Leaf smut. Each class of rice plant disease consists of 40 images per class. The sample data can be shown in Figure 4. The dataset that has been collected is then processed. Image processing is essential in improving the quality of processed images so that the CNN model is more optimal for predicting rice plant diseases. Image processing techniques are resizing, converting to grayscale, and data augmentation. Resize image serves to homogenize the size of an image to reduce computation time. This research uses an image size of 200 * 200.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

(a)

(b)

(c)

Fig 4. Sample Dataset; (a). Bacterial Leaf Blight, (b). Brown Spot, (c). Leaf Smut

The next step is to convert the RGB image to grayscale. Changing an RGB image to a grayscale image makes the execution time and complexity lower, as the complexity level can affect the model's performance in image recognition. The further process is data augmentation. The amount of data obtained is still tiny, so the data does not vary too much; as a result, the CNN method is not optimal in recognizing the data and affects the level of accuracy. Therefore, data augmentation is needed to add more data so that it is more varied so that the CNN method is more optimal in recognizing rice disease data. The data augmentation process uses several techniques so that the augmented data is not 100% similar to the original data. The augmentation techniques used are Rescale, Shear, Zoom, Rotation, and Flip Horizontal. The amount of data before and

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

after data augmentation is shown in Table 1. In contrast, the sample data generated by the augmentation process is shown in Figure 5.

Table 1. Total Data on Rice Plant Disease Types Before and After Augmentation

No	Class of Rice Disease	Data Original	Data Augmentation
1	Bacterial leaf blight	40	1025
2	Brown spot	40	1039
3	Leaf smut	40	982

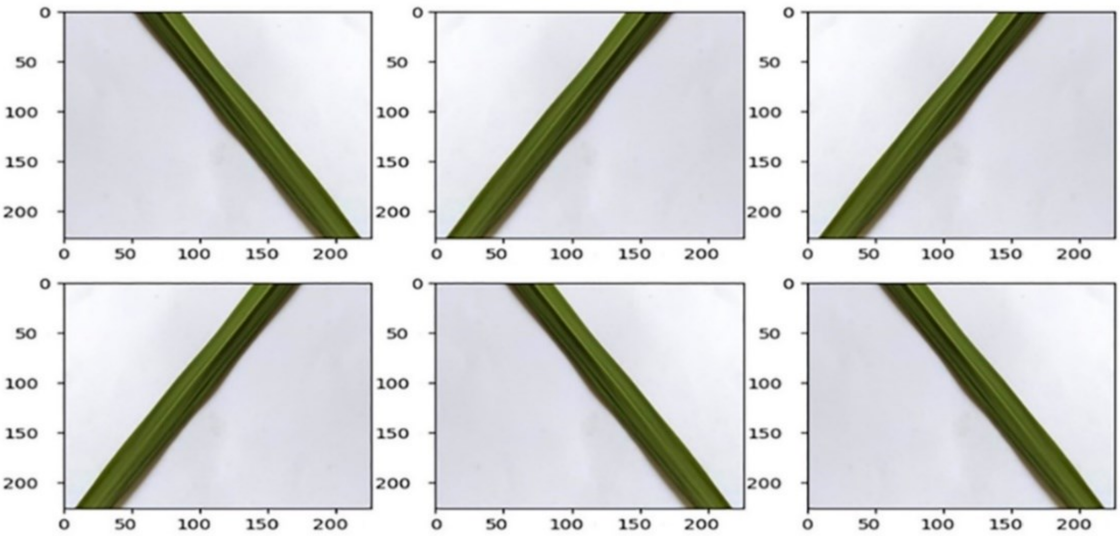


Fig 5. Data Augmentation Process Result

Image data of rice plant leaves that have been carried out data augmentation is then detected using the CNN method. The CNN method can classify the type of rice plant disease by learning the training data pattern. The more training data given, the smarter the CNN method recognizes the data pattern. The CNN method, learned from the training data, will then be tested with testing data to determine the success rate in detecting the type of rice plant disease. This research uses 80% training data (2437) and 20% testing data (609). Then, the CNN method process in detecting the type of rice plant disease uses the architecture as in Figure 6.

INTENSIF, Vol.8 No.1 February 2024
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

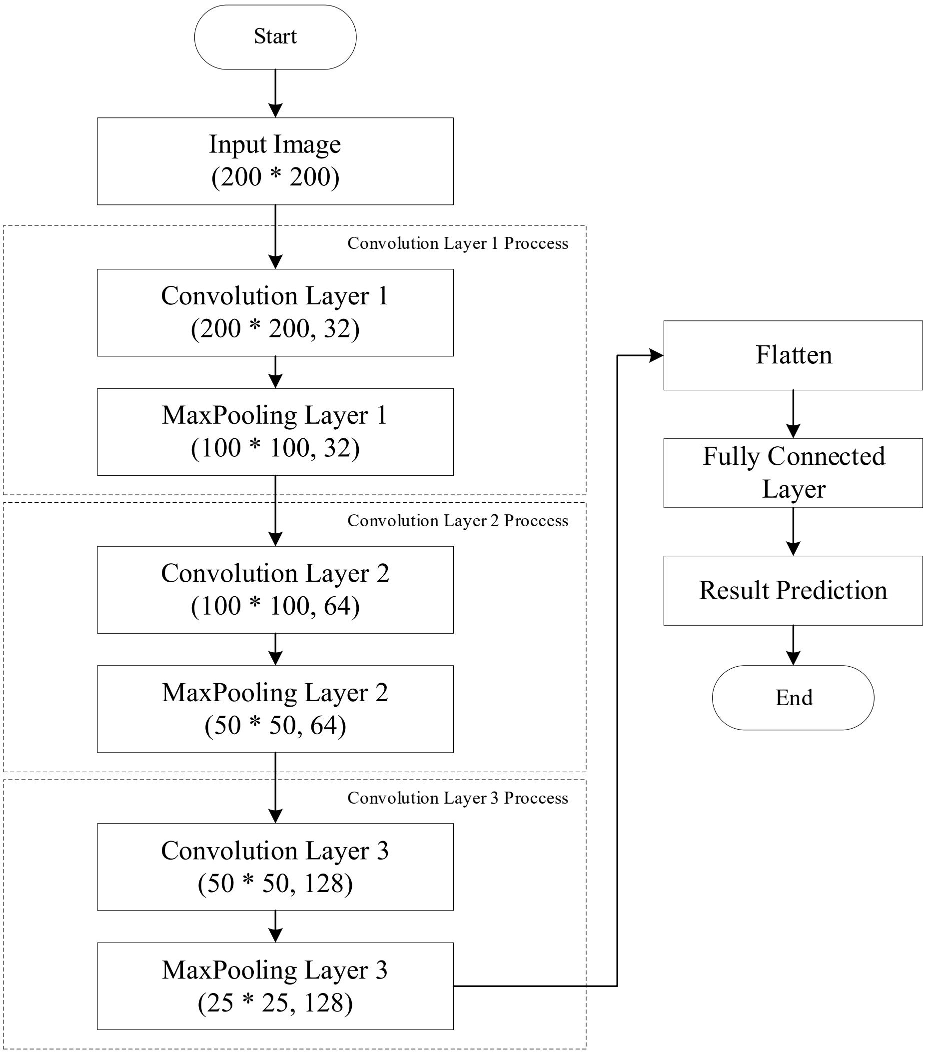


Fig 6. CNN Model Architecture

Based on Figure 6, the CNN method process in detecting the type of rice plant disease starts from the input image of rice leaves measuring 200 * 200. Then convolution is carried out three times for feature extraction in the image and reducing features using the Max Pooling method three times so that the final size of the feature is 25 * 25. Furthermore, the Flatten process is carried out to change the feature size of 25 * 25 into a 1-dimensional vector used as input. In Figure 7, the CNN method correctly classified the Bacterial leaf blight class as much as 204 out of 205, the Brown spot class was correctly classified as 207 data, and the Leaf smut class was successfully classified as 195 data out of 196. Based on Figures 8.a and 6.b regarding the accuracy and loss graph of the CNN method for 100 epochs. The greater the number of epochs, the accuracy performance increases. Conversely, the greater the number of epochs, the lower the loss rate. The difference in accuracy and loss values for each training data and testing data is not too far, this indicates that overfitting and underfitting does not occur.

INTENSIF, Vol.8 No.1 February 2024
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

Thus, the CNN method correctly classified as much as 606 data from a total of 608 data so that the resulting accuracy was 99.7%. The CNN method is highly accurate inseparable from the data augmentation process. Researchers [34][35][36][37] state that the use of data augmentation can improve the performance of the method used. The comparison of the findings of this research with previous research is shown in Table 2. In Table 2, the proposed method (CNN with Data Augmentation) has better accuracy than previous research in detecting rice leaf disease types. The proposed method is better than the previous research method because there is an additional data augmentation process, so the accuracy of the CNN method is very high like figure 8.

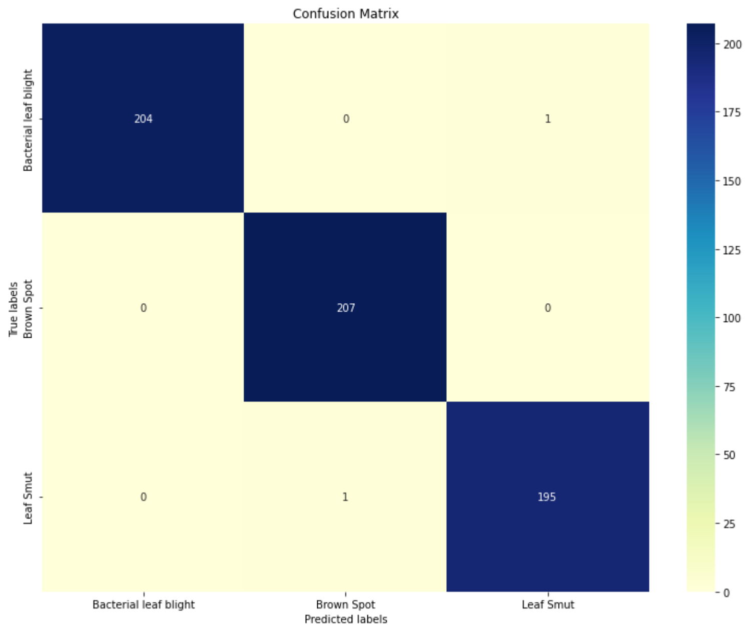
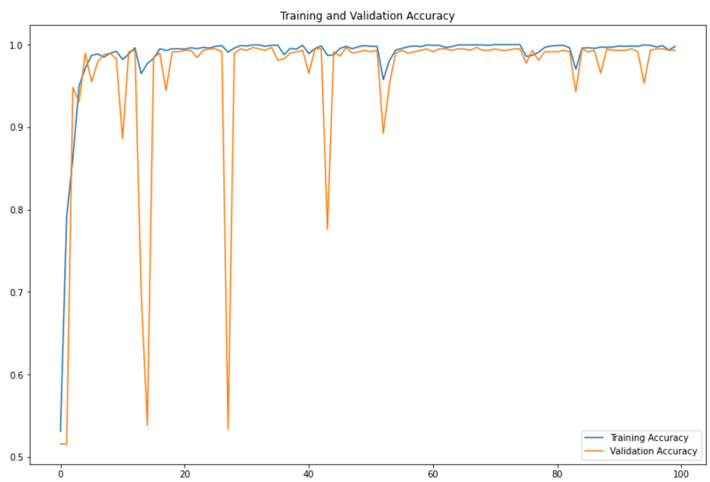
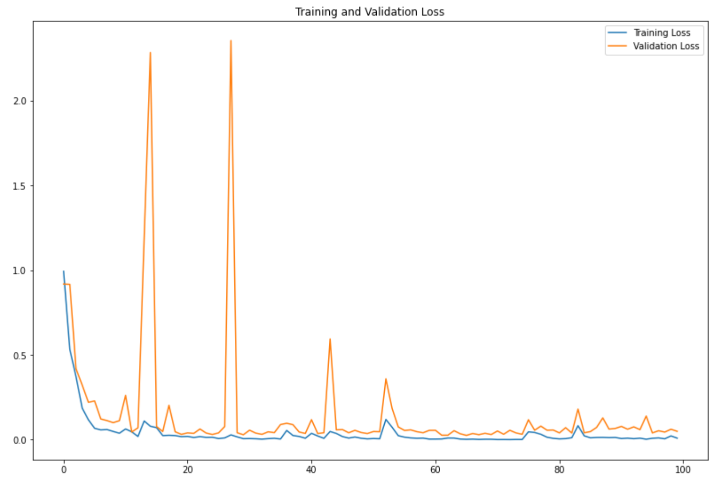


Fig 7. Confusion Matrix Results of CNN Method on Rice Plant Disease Detection



(a)

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168



(b)

Fig 8. (a) Accuracy Graph of CNN Method with 100 Epochs, (b) Loss Graph of CNN Method with 100 Epochs

Table 2. Comparison Result of this Research with Previous Research

No	Researcher(s)	Methods	Case Study	Accuracy
1	Khoiruddin, et al [38]	CNN	Rice Plant Diseases	98%
2	Primatua, et.al [39]	DenseNet 201		83%
3	Tejaswini, et.al [22]	CNN		78%
4	Ruoling, et.al [23]	VGGNet		92%
5	Lilis, et.al [18]	Gray Level Co-occurrence Matrix (GLCM)		90%
6	Dwijayana, et.al [40]	MobileNetV2		94,6%
7	**Proposed method**	**Augmented Data with CNN**		**99,7%**

IV. CONCLUSION

The CNN method has been successfully implemented in detecting the type of rice leaf disease with a division of 80% training data and 20% testing data. The CNN method obtained an accuracy rate of 99.7% for detecting leaf disease types of rice plants. The CNN method can obtain very high accuracy in detecting rice leaf diseases because there is a data augmentation process, so the data is more diverse. Future research can perform parameter tuning to obtain optimal parameters to improve the performance of the CNN method. Not only that, this research does not resolve the noise in the image, so future focused on resolving the noise.

Author Contributions: *Hairani*: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision. *Triyanna Widiyaningtyas*: Investigation, Data Curation, Writing.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: The data cannot be openly shared for the protection of study participant privacy.

Informed Consent: There were no human subjects.

Animal Subjects: There were no animal subjects.

ORCID:
Hairani: http://orcid.org/0000-0002-6756-5896
Triyanna Widiyaningtyas: http://orcid.org/0000-0001-6104-6692

REFERENCES

[1] J. L. Sukowati, "The Future of Indonesia and Global Agriculture: Rice Comsumption and Agricultural Modernization," Jurnal Litbang Sukowati, vol. 4, no. 1, pp. 57–64, 2020, doi: 10.32630/sukowati.v4i1.132.

[2] M. Murdiyana and M. Mulyana, "Analisis kebijakan pengentasan kemiskinan di indonesia," Jurnal Politik Pemerintahan, vol. 10, no. 1, pp. 73–96, 2019.

[3] E. Erythrina et al., "Assessing Opportunities to Increase Yield and Profit in Rainfed Lowland Rice Systems in Indonesia," Agronomy, vol. 11, no. April, pp. 1–15, 2021, doi: https://doi.org/ 10.3390/agronomy11040777.

[4] M. F. Umakamea et al., "Kerusakan Lima Varietas Padi Akibat Serangan Hama Penggerek Batang di Desa Savanajaya, Kecamatan Waeapo, Kabupaten Buru Damages," Jurnal Budidaya Pertanian, vol. 16, no. 2, pp. 180–186, 2020, doi: 10.30598/jbdp.2020.16.2.180.

[5] X. Sun, J. Lyu, and C. Ge, "Knowledge and Farmers' Adoption of Green Production Technologies: An Empirical Study on IPM Adoption Intention in Major Indica-Rice-Producing Areas in the Anhui Province of China," Int J Environ Res Public Health, vol. 19, no. 21, pp. 1–16, Nov. 2022, doi: 10.3390/ijerph192114292.

[6] Md. M. Hasan et al., "Enhancing Rice Crop Management: Disease Classification Using Convolutional Neural Networks and Mobile Application Integration," Agriculture, vol. 13, no. 8, pp. 1–17, Aug. 2023, doi: 10.3390/agriculture13081549.

[7] N. Zeng, G. Gong, G. Zhou, and C. Hu, "An Accurate Classification of Rice Diseases Based on ICAI-V4," Plants, vol. 12, no. 11, pp. 1–28, 2023, doi: 10.3390/plants12112225.

[8] A. Nigam, A. K. Tiwari, and A. Pandey, "Paddy leaf diseases recognition and classification using PCA and BFO-DNN algorithm by image processing," Mater Today Proc, vol. 33, pp. 4856–4862, 2020, doi: 10.1016/j.matpr.2020.08.397.

[9] M. Keskar and D. D. Maktedar, "Hybrid deep-spatio textural feature model for medicinal plant disease classification," Indonesian Journal of Electrical Engineering and Computer Science, vol. 30, no. 1, p. 356, 2023, doi: 10.11591/ijeecs.v30.i1.pp356-365.

[10] Y. Ramdhani, D. F. Apra, and D. P. Alamsyah, "Feature selection optimization based on genetic algorithm for support vector classification varieties of raisin," Indonesian Journal of Electrical Engineering and Computer Science, vol. 30, no. 1, p. 192, 2023, doi: 10.11591/ijeecs.v30.i1.pp192-199.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

[11] G. Latif, S. E. Abdelhamid, R. E. Mallouhy, J. Alghazo, and Z. A. Kazimi, “Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model,” Plants, vol. 11, no. 17, 2022, doi: 10.3390/plants11172230.

[12] Meeradevi and M. R. Mundada, “Hybrid Decision Support System Framework for Leaf Image Analysis to Improve Crop Productivity,” International Journal of Advanced Computer Science and Applications, vol. 12, no. 9, pp. 381–387, 2021, doi: 10.14569/IJACSA.2021.0120943.

[13] S. F. Wani, A. Ashraf, and A. Sophian, “Applied Research and Smart Technology,” Applied Research and Smart Technology, vol. 3, no. 1, pp. 38–48, 2022.

[14] R. R. Atole and D. Park, “A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies,” International Journal of Advanced Computer Science and Applications, vol. 9, no. 1, pp. 67–70, 2018, doi: 10.14569/IJACSA.2018.090109.

[15] M. A. Azim, M. K. Islam, M. M. Rahman, and F. Jahan, “An effective feature extraction method for rice leaf disease classification,” Telkomnika (Telecommunication Computing Electronics and Control), vol. 19, no. 2, pp. 463–470, 2021, doi: 10.12928/TELKOMNIKA.v19i2.16488.

[16] A. D. Saputra, D. Hindarto, and H. Santoso, “Disease Classification on Rice Leaves using DenseNet121, DenseNet169, DenseNet201,” Sinkron, vol. 8, no. 1, pp. 48–55, 2023, doi: 10.33395/sinkron.v8i1.11906.

[17] K. N, L. V. Narasimha Prasad, C. S. Pavan Kumar, B. Subedi, H. B. Abraha, and S. V E, “Rice leaf diseases prediction using deep neural networks with transfer learning,” Environ Res, vol. 198, p. 111275, Jul. 2021, doi: 10.1016/j.envres.2021.111275.

[18] L. Indrayani and R. Wirawan, “Implementation of Gray Level Coocurence Matrix on the Leaves of Rice Crops,” Jurnal Ilmiah Teknik Elektro Komputer dan Informatika, vol. 6, no. 1, p. 1, 2020, doi: 10.26555/jiteki.v16i1.16630.

[19] Faza Adhzima, Yandra Arkeman, and Irman Hermadi, “The Clustering Rice Plant Diseases Using Fuzzy C-Means and Genetic Algorithm,” Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 6, no. 2, pp. 240–245, 2022, doi: 10.29207/resti.v6i2.3912.

[20] K. Kethineni and G. Pradeepini, “Identification of Leaf Disease Using Machine Learning Algorithm for Improving the Agricultural System,” Journal of Advances in Information Technology, 2023, doi: 10.12720/jait.14.1.122-129.

[21] A. A. J. V. Priyangka and I. M. S. Kumara, “Classification Of Rice Plant Diseases Using the Convolutional Neural Network Method,” Lontar Komputer : Jurnal Ilmiah Teknologi Informasi, vol. 12, no. 2, p. 123, 2021, doi: 10.24843/lkjiti.2021.v12.i02.p06.

[22] P. Tejaswini, P. Singh, M. Ramchandani, Y. K. Rathore, and R. R. Janghel, “Rice Leaf Disease Classification Using Cnn,” in IOP Conference Series: Earth and Environmental Science, 2022, pp. 1–14. doi: 10.1088/1755-1315/1032/1/012017.

[23] R. Deng et al., “Automatic Diagnosis of Rice Diseases Using Deep Learning,” Front Plant Sci, vol. 12, no. August, Aug. 2021, doi: 10.3389/fpls.2021.701038.

[24] K. Kishore Kumar and E. Kannan, “An Efficient Deep Neural Network for Disease Detection in Rice Plant Using XGBOOST Ensemble Learning Framework,” International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 3, pp. 116–128, 2022.

[25] N. Senan, M. Aamir, R. Ibrahim, N. S. A. M. Taujuddin, and W. H. N. W. Muda, “An efficient convolutional neural network for paddy leaf disease and pest classification,” International Journal of Advanced Computer Science and Applications, vol. 11, no. 7, pp. 116–122, 2020, doi: 10.14569/IJACSA.2020.0110716.

[26] G. K. V. L. Udayananda, C. Shyalika, and P. P. N. V. Kumara, “Rice plant disease diagnosing using machine learning techniques: a comprehensive review,” SN Appl Sci, vol. 4, no. 11, 2022, doi: 10.1007/s42452-022-05194-7.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v8i1.21168

[27] A. R. Khan, I. Abunadi, B. AlGhofaily, H. Ali, and T. Saba, "Automatic Diagnosis of Rice Leaves Diseases Using Hybrid Deep Learning Model," Journal of Advances in Information Technology, vol. 14, no. 3, pp. 418–425, 2023, doi: 10.12720/jait.14.3.418-425.

[28] H. Yang, X. Deng, H. Shen, Q. Lei, S. Zhang, and N. Liu, "Disease Detection and Identification of Rice Leaf Based on Improved Detection Transformer," Agriculture, vol. 13, pp. 1–17, 2023, doi: https://doi.org/10.3390/ agriculture13071361.

[29] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance," in Applications in Electronics Pervading Industry, Environment and Society, S. Saponara and A. De Gloria, Eds., Cham: Springer International Publishing, 2022, pp. 10–17.

[30] J. Guo, J. Ma, Á. F. García-Fernández, Y. Zhang, and H. Liang, "A survey on image enhancement for Low-light images," Heliyon, vol. 9, no. 4. Elsevier Ltd, pp. 1–26, Apr. 01, 2023. doi: 10.1016/j.heliyon.2023.e14558.

[31] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," Array, vol. 16. Elsevier B.V., Dec. 01, 2022. doi: 10.1016/j.array.2022.100258.

[32] H. Hairani, A. Anggrawan, and D. Priyanto, "Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link," International Journal on Informatics Visualization, vol. 7, no. 1, pp. 258–264, 2023.

[33] S. Rezvani and X. Wang, "A broad review on class imbalance learning techniques," Appl Soft Comput, vol. 143, p. 110415, 2023, doi: 10.1016/j.asoc.2023.110415.

[34] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," Artificial Intelligence in Agriculture, vol. 6, pp. 23–33, 2022, doi: 10.1016/j.aiia.2021.12.002.

[35] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J Big Data, vol. 6, no. 1, pp. 1–48, 2019, doi: 10.1186/s40537-019-0197-0.

[36] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," Global Transitions Proceedings, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.

[37] N. Marastoni, R. Giacobazzi, and M. Dalla Preda, "Data augmentation and transfer learning to classify malware images in a deep learning context," Journal of Computer Virology and Hacking Techniques, vol. 17, no. 4, pp. 279–297, Dec. 2021, doi: 10.1007/s11416-021-00381-3.

[38] M. Khoiruddin, A. Junaidi, and W. A. Saputra, "Klasifikasi Penyakit Daun Padi Menggunakan Convolutional Neural Network," Journal of Dinda, vol. 2, no. 1, pp. 37–45, 2022.

[39] P. Sitompul, H. Okprana, and A. Prasetio, "Identification of Rice Plant Diseases Through Leaf Image Using DenseNet 201," JOMLAI: Journal of Machine Learning and Artificial Intelligence, vol. 1, no. 2, pp. 143–150, 2022, doi: 10.55123/jomlai.v1i2.889.

[40] I. G. D. Dwijayana and I. G. A. Wibawa, "Implementasi Transfer Learning Dalam Klasifikasi Penyakit Pada Daun Teh Menggunakan MobileNetV2," Jurnal Nasional Teknologi Informasi dan Aplikasinya, vol. 1, no. 1, pp. 379–387, 2022.