Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 1 Iss. 2 January 2024, pp: 71-77
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease

Delvi Hastari[1*], Salsa Winanda[2], Aditya Rezky Pratama[3], Nana Nurhaliza[4], Ella Silvana Ginting[5]

[1,2,3,4]Department of Information System, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[5]Faculty of Technology Management and Business, University Tun Hussein Onn Malaysia, Malaysia

E-Mail: [1]delvihastari19@gmail.com, [2]sals34549@gmail.com, [3]adityarezky121@gmail.com, [4]nananurhaliza@gmail.com, [5]ellaginting86@gmail.com

Received Aug 13th 2023; Revised Dec 3rd 2023; Accepted Dec 30th 2023
Corresponding Author: Delvi Hastari

Abstract

Rice is the most important crop in global food security and socioeconomic stability. A part of the world's population makes rice a food requirement but the problem is found that all rice varieties suffer from several diseases and pests. Therefore, it is necessary to ensure the quality of healthy and proper rice growth by detecting diseases present in rice plants and treatment of affected plants. In this study, the Convolutional Neural Network (CNN) algorithm was applied in classifying diseases on the leaves of rice plants by experimenting with several parameters and architecture to get the best accuracy. This study was conducted image classification of rice plant disease using CNN architecture ResNet-50V2 with data using preprocessing Augmentation. The test was conducted with three optimizers such as SGD, Adam, and RMSprop by combining various parameters, namely epoch, batch size, learning rate, and SGD and RMSprop optimizers. Division of image data with 70:30 ratio of training data and test data; 80:20; 90:10. From these results, it was found that Adam was the best optimizer in the 80:20 data division in this study with an accuracy level of 0.9992, followed by the SGD optimizer with an accuracy level of 0.9983, while the RMSProp optimizer was ranked third with an accuracy level of 0.9978.

Keyword: Convolutional Neural Network, Images Classification, Resnet-50 V2, Rice Plant Disease

1. INTRODUCTION

Rice is the most important crop in global food security and socioeconomic stability. Most of the world's population makes rice a food requirement but the problem is found that all rice varieties suffer from several diseases and pests [1]. Therefore, it is necessary to ensure the quality of healthy and proper rice growth by detecting diseases present in rice plants and treatment of affected plants [2]. Treatment of rice plants must be in accordance with the type of disease that exists, but the problem that occurs is that the process of classifying diseases from symptoms is still manually, which is laborious and errors can occur in classification [3].

Research on diseases in rice plants has previously been carried out by Kawcher and colleagues by presenting a rice leaf disease detection system using a machine learning approach. The rice plant disease dataset was trained with a variety of different machine learning algorithms including KNN (K-Nearest Neighbour), J48 (Decision Tree), Naive Bayes and Logistic Regression [2]. Then the study [4] which presents an image-based approach to classifying rice plant diseases only uses color features. The result was the highest classification accuracy of 94.65% using the Support Vector Machine (SVM) Algorithm.

Many algorithms can be used to classify. Research [5] compared Support Vector Machine (SVM), Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) algorithms in classifying vegetation species using Hyperspectral Thermal Infrared Dataset. The results show that the CNN algorithm has almost perfect accuracy, reaching 99% and being the highest compared to other algorithms. Thus, the CNN algorithm is very good to be used in classifying plant leaf diseases, especially rice. Research by [6] conducted experiments by introducing the use of deep learning with the latest CNN algorithm related to plant leaf disease classification. The use of deep learning can be done in detecting and classifying plant diseases [7]. The research proves that this deep learning method provides a solution to the problem of traditional machine methods.

Some researchers have used the CNN algorithm in classifying diseases, such as studies conducted by [8] to classify breast cancer using the CNN algorithm combined with the Bit-Plane feature, showing that the proposed method can greatly improve recognition rates and classification performance. In study [9] or the classification of diffuse liver disease based on ultrasound images with multimodal features, the method used can increase overall classification accuracy by 5.4%. Furthermore, research [10] conducted to classify disease symptoms in maize plants by combining CNN with Bidirectional Long Short-Term Memory model showed good results of 99.02%. Thus, the selection of this research is based on the success of the CNN method in previous studies and its potential in providing an accurate and effective solution to the problem of rice leaf disease classification.

The CNN-based plant leaf disease recognition model largely has limitations on shooting conditions and background. The problems with CNN-based classification are sometimes limited to generalizations such as accuracy dropping dramatically when new data is entered into a dataset. In addition, it is necessary to consider the effectiveness of tissue parameters for better recognition of foliar diseases [1]. Therefore, this study applied the CNN algorithm in classifying diseases on rice leaves with experiments on several parameters and architecture to get the best accuracy.

2. MATERIAL AND METHOD

2.1 Rice Plant Disease

Rice plant disease is a condition that affects the growth and health of the rice plant, one of the most important food crops in the world. Diseases of rice plants can be caused by insect attacks of pests, pathogens, nutritional deficiencies and environmental conditions [11]. There are various diseases that can attack rice plants, such as bacterialblight, blast, brown spot, and tungro.

1. Brown Spot
   The disease is caused by the fungus Cochliobolus miyabeanus. Symptoms of brown spot are visible on leaves, where oval or elongated brown spots with clear edges are formed. The spots may then enlarge and merge to form large, dry-brown lesions. Brown spot disease tends to develop in humid and warm conditions [12].

2. Bacterial Blight
   Bacterial blight, or bacterial wilt, is a rice plant disease caused by the bacterium Xanthomonas oryzae pv. oryzae. Symptoms include the appearance of grayish-green spots on the leaves which then turn dark brown. Such lesions often have a long, transverse shape on leaves, stems, panicles and roots. Infected plants can also experience poor germination, stunted growth, and even death [13].

3. Blast
   This disease is caused by the fungus Pyricularia oryzae and can attack various parts of the rice plant, including leaves, stems, and panicles. Blast usually appears in humid environmental conditions, especially during rapid rice growth. Some of the symptoms that appear on the leaves are rhomboid-shaped brown spots and extending in the direction of the leaf veins, brown rickshaw edges with a grayish-white middle, and gray flecks [14].

4. Tungro
   This disease is caused by a viral complex consisting of Rice tungro bacilliform virus (RTBV) and Rice tungro spherical virus (RTSV). Tungro disease is transmitted through vector insects, namely brown leafhoppers, which suck the juice of infected plants and spread it to healthy rice plants. Symptoms of tungro in rice plants include stunted growth, yellowing of leaves, wavy leaves, and disturbed panicle formation [15].

The research stages for the classification of rice plant disease images in this study are shown in Flowchart or Figure 1.

2.2 Data Collection

The data collected is image data of leaf disease in rice plants. This secondary data was obtained from the Kaggle site (https://www.kaggle.com/datasets/shareef0612/riceleaf-dataset). The data used amounted to 5932 images consisting of 1584 Bacterial blight images, 1440 Blast images, 1600 Brown spot images and 1308 Tungro images. Each class will share images for training and testing to train and test models. The data will be divided into a training subset, a validation subset, and a testing subset.

2.3 Data Sharing

Hold-out techniques are used to divide data, into training data and testing data. Hold-out is a method that will provide a certain amount of data to be used as training data and the rest as testing data. In this study, data distribution was carried out using three different ratios, namely 70:30, 80:20, and 90:10.

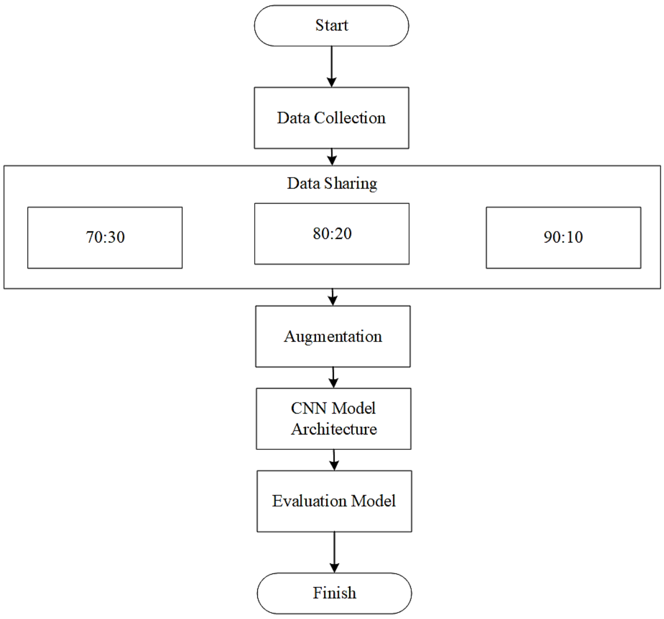

Figure 1. Research Methodology

2.4 Augmentation

Augmentation techniques are intended to enrich the image without losing its substance through the process of image modification based on certain parameters. In addition, performing augmentation aims to improve the accuracy of the CNN model. The augmentation process is carried out by utilizing the Image Data Generator function from Tensorflow. The augmentation parameters used include rotation range, rescale, shear range, zoom range, horizontal flip and fill mode.

2.5 CNN Model Architecture

The primary principle behind Convolutional Neural Network is to extract local features from high layer input and transmit them to lower layers for more complicated features [16]. This type of network is primarily designed for processing topological data, such as grids or images [17]. CNNs typically comprise an input layer, a convolutional layer, pooling, and a fully connected layer [18]. CNN is a neural network with some convolutional layers (and some other layers). Convolutional layers compute the feature map using kernel input and activation as stated in Equation 1 below [14].

$$m_i^l = f\left(x_j^l\right) \tag{1}$$

Where m_i^l is the number of feature maps to calculate, identify first layer. The function is used for activation (relu, sigmoid). The activation of the j channel in the given layer 1 is provided by equation 2 [14].

$$X_j^l = \sum_{i \in a}^{I} z_i^{l-1} * c_{ij}^i + p_j^l \tag{2}$$

c_{ij}^i is the convolutional filter matrix, and p_j^l denotes the offset at first layer. The convolutional layer operates by leveraging local information correlations within the image to extract distinctive features [19]. Through sampling, the pooling layer selects relevant features from the higher-level feature map while simultaneously establishing invariant models for translation, rotation, and scaling [20]. In this study, the ResNet-50 V2 architecture was used as a model of Convolutional Neural Network (CNN) which is an advanced convolutional network that is easier to train than other deep convolutional neural networks, offering better accuracy and faster convergence [21]. In this ResNet-50 V2 architecture, there are several types of layers used to process image data on datasets rice leaf gradually and eventually generate predictions of the desired class such as the use of layers Dense, Flatten and Dropout.

2.6 Optimizer

A deep learning optimization method of choice determines the training pace and eventual predictive performance of model [22]. Adam, RMSProp and SGD optimization methods are used to reduce errors throughout the training process. Adam optimizer features a pretty efficient computing procedure that does not use gradient scales and can handle massive amounts of data or parameters [21]. RMSprop optimizer is a mini-batch learning adaptive variant of the rprop method [17]. The SGD excels in terms of convergence speed [23].

2.7 Evaluation Model

Evaluation models are very important in verifying classification performance on modeling [24]. For classification, general accuracy is used to evaluate the developed model. Accuracy is the percentage of compounds that are correctly classified in the total number of compounds, as shown in Equation 3 [25].
The formula is written clearly using an equation with an index like the following example:

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{3}$$

3. RESULTS AND DISCUSSION

The experimental process in this study used Google Colab in running the model. The programming language used is python and the use of Keras and Tensorflow libraries. The parameters used are learning rate with a value of 0.0001, dense with a value of 512 and 256, batch size with a value of 64, and Optimizer Adam, RMSProp and SGD and uses the ResNet-50 V2 architecture. Data ratios are 70:30, 80:20, 90:10. The Table 1 shows the test results in this study.

Table 1. Testing Results

Split Data		Batch Size	Learning Rate	Epoch	Optimizer	Accuracy
Training	Testing					
70	30	64	0.0001	10	Adam	0.9927
				9	RMSProp	0.9944
				100	SGD	0.9927
80	20	64	0.0001	15	Adam	0.9992
				12	RMSProp	0.9978
				100	SGD	0.9941
90	10	64	0.0001	10	Adam	0.9933
				11	RMSProp	0.9955
				6	SGD	0.9983

Table 1 shows the percentage results of 9 experiments with optimizer combinations that have been performed. From the table, you can see the difference in the percentage accuracy value depending on the optimizer used.

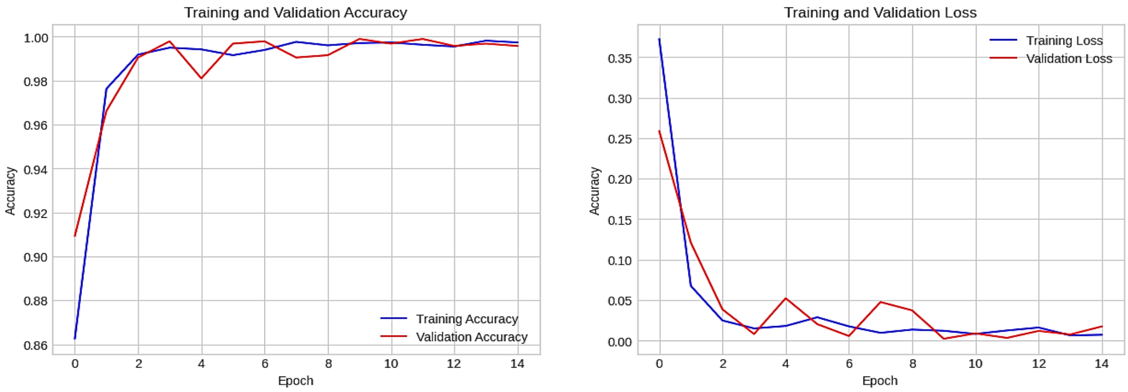

Figure 2. Graphic Accuracy and Loss Adam Highest Performance

Figure 2 is a graph of the accuracy and loss model in 80:20 data division, where the graph is the best performing graph produced by ResNet-50V2 with Adam optimizer. The results were obtained with parameters in the form of batch size 64, learning rate 0.0001 and epoch 15 with accuracy of 0.9992. The accuracy value is good because the value is close to number 1. Figure 2 also shows a graph of the loss value gained, rising at a certain point and falling to epoch 15 which is 0.0035. The loss value is also good because the value is close to 0.

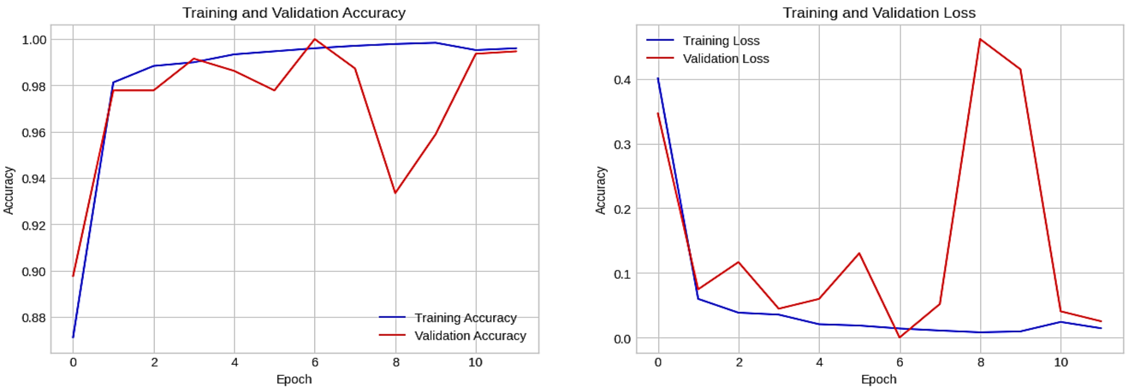

Figure 3. Graphic Accuracy and Loss RMSProp Highest Performance

Figure 3 is a graph of accuracy and loss models in 80:20 data division, where the graph is the best performing graph produced by ResNet-50V2 with RMSprop optimizer. The results were obtained with parameters in the form of batch size 64, learning rate 0.0001 and epoch 12 with accuracy of 0.9978. The accuracy value is good because the value is close to number 1. Figure 3 also shows a graph of the loss value gained, rising at a certain point and falling to epoch 12 of 0.0056. The loss value is also good because the value is close to 0.

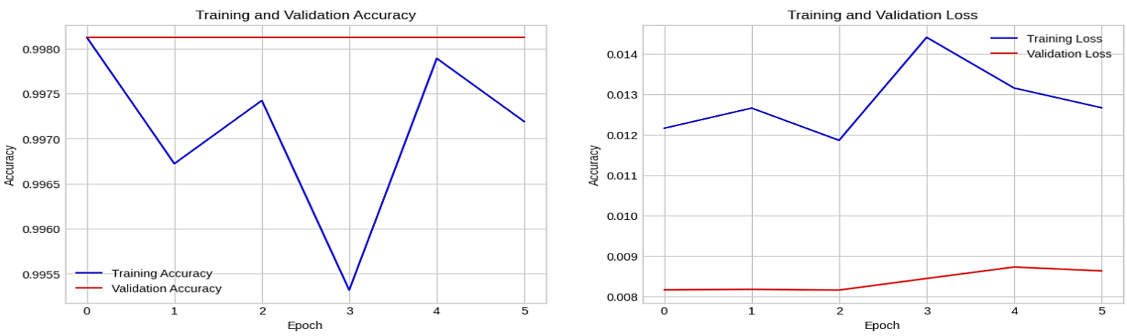

Figure 4. Graphic Accuracy and Loss SGD Highest Performance

Figure 4 is a graph of accuracy and loss models in 90:10 data division, where the graph is the best performing graph produced by ResNet-50V2 with SGD optimizer. The results were obtained with parameters in the form of batch size 64, learning rate 0.0001 and epoch 6 with accuracy of 0.9983. The accuracy value is good because the value is close to number 1. Figure 4 also shows a graph of the loss value gained, rising at a certain point and falling to epoch 6 which is 0.0039. The loss value is also good because the value is close to 0.

Table 2. Performance Optimizer Overview

Optimizer	Split Data	Epoch	Accuracy	Average
Adam	80:20	15	0.9992	83,00
RMSProp	80:20	12	0.9978	80,92
SGD	90:10	6	0.9983	82,53

Table 2 summarizes the best accuracy models of each optimizer. Adam's optimizer with a batch size of 64, learning rate of 0.0001 and epoch of 15 has the best performance compared to RMSProp optimizer with the same data share of 80:20. While the SGD optimizer with a batch size value of 64, learning rate of 0.0001 and epoch 6 with a data share of 90:10, has the best performance compared to the RMSProp optimizer.

4. CONCLUSION

In this study, image classification of rice plant diseases was carried out using CNN architecture ResNet-50V2 with data using preprocessing Augmentation. The test was applied hold-out data sharing techniques as well as optimizers Adam, RMSProp and SGD by combining various parameters, namely epoch, batch size and learning rate. The experimental results showed that the 80:20 hold-out data separation technique and Adam's optimizer had the best accuracy result of 0.9992, followed by the SGD optimizer with an accuracy level of 0.9983, while the RMSProp optimizer was ranked third with an accuracy level of 0.9978. The results showed that Data sharing and optimizers are also important things that can affect the level of accuracy.

REFERENCES

[1] S. M. M. Hossain *et al.*, *Rice Leaf Diseases Recognition Using Convolutional Neural Networks*, vol. 12447 LNAI. Springer International Publishing, 2020. doi: 10.1007/978-3-030-65390-3_23.

[2] K. Ahmed, T. R. Shahidi, S. M. I. Alam, and S. Momen, "Rice Leaf Disease Detection Using Machine Learning Techniques," *2019 Int. Conf. Sustain. Technol. Ind. 4.0*, pp. 1–5, 2019, doi: 10.1109/STI47673.2019.9068096.

[3] C. R. Rahman *et al.*, "Identification and recognition of rice diseases and pests using convolutional neural networks," *Biosyst. Eng.*, vol. 194, pp. 112–120, 2020, doi: 10.1016/j.biosystemseng.2020.03.020.

[4] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: a machine learning paradigm," *J. Plant Pathol.*, vol. 103, no. 2021, pp. 17–26, 2020, doi: 10.1007/s42161-020-00683-3.

[5] M. ul Hasan, S. Ullah, M. J. Khan, and K. Khurshid, "Comparative analysis of SVM, ann and cnn for classifying vegetation species using hyperspectral thermal infrared data," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 42, no. 2/W13, pp. 1861–1868, 2019, doi: 10.5194/isprs-archives-XLII-2-W13-1861-2019.

[6] P. Padma and J. Ahn, "Guest satisfaction & dissatisfaction in luxury hotels: An application of big data," *Int. J. Hosp. Manag.*, vol. 84, pp. 1–8, 2020, doi: 10.1016/j.ijhm.2019.102318.

[7] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 707, pp. 1–18, 2021, doi: 10.3390/agriculture11080707.

[8] G. Chen, Y. Chen, Z. Yuan, X. Lu, X. Zhu, and W. Li, "Breast Cancer Image Classification based on CNN and Bit-Plane slicing Guoming," *2019 Int. Conf. Med. Imaging Phys. Eng. (CMIPE)*, pp. 7–10, 2019.

[9] L. Dandan, M. Huanhuan, L. Xiang, J. Yu, J. Jing, and S. Yi, "Classification of diffuse liver diseases based on ultrasound images with multimodal features," *2019 IEEE Int. Instrum. Meas. Technol. Conf.*, pp. 1–5, 2019, doi: 10.1109/I2MTC.2019.8827174.

[10] M. J. Hasan, M. S. Alom, U. F. Dina, and M. H. Moon, "Maize Diseases Image Identification and Classification by Combining CNN with Bi-Directional Long Short-Term Memory Model," *2020 IEEE Reg. 10 Symp.*, pp. 1804–1807, 2020, doi: 10.1109/TENSYMP50017.2020.9230796.

[11] A. Rifa'I and D. Mahdiana, "Image processing for diagnosis rice plant diseases using the fuzzy system," *2020 Int. Conf. Comput. Sci. Its Appl. Agric.*, pp. 1–5, 2020, doi: 10.1109/ICOSICA49951.2020.9243274.

[12] S. Ghosal and K. Sarkar, "Rice Leaf Diseases Classification Using CNN with Transfer Learning," *Proc. 2020 IEEE Calcutta Conf.*, pp. 230–236, 2020, doi: 10.1109/CALCON49167.2020.9106423.

[13] N. Manohar and K. J. Gowda, "Image Processing System based Identification and Classification of Leaf Disease: A Case Study on Paddy Leaf," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC 2020)*, pp. 451–457, 2020, doi: 10.1109/ICESC48915.2020.9155607.

[14] A. K and A. S. Singh, "Detection of Paddy Crops Diseases and Early Diagnosis Using Faster Regional Convolutional Neural Networks," *2021 Int. Conf. Adv. Comput. Innov. Technol. Eng.*, pp. 898–902, 2021, doi: 10.1109/ICACITE51222.2021.9404759.

[15] D. S. K. RAO, D. M. KRISHNA, and D. B. RAMAKRISHNA, "SMART AILMENT IDENTIFICATION SYSTEM FOR PADDY CROP USING MACHINE LEARNING," *Int. J. Innov. Eng. Manag. Res.*, vol. 09, no. 03, pp. 96–100, 2020.

[16] S. T. Rizaldi, Mustakim, I. Permana, and M. Afdal, "Image Enhancement on Deep Learning Algorithm for COVID-19 Lung X-Ray Classification," *Proceeding - 2022 Int. Symp. Inf. Technol. Digit. Innov. Technol. Innov. Dur. Pandemic, ISITDI 2022*, pp. 11–15, 2022, doi: 10.1109/ISITDI55734.2022.9944501.

[17] M. A. Islam, S. I. Rashid, N. U. I. Hossain, R. Fleming, and A. Sokolov, "An integrated convolutional neural network and sorting algorithm for image classification for efficient flood disaster management," *Decis. Anal. J.*, vol. 7, p. 100225, 2023, doi: 10.1016/j.dajour.2023.100225.

[18] H. Deng, W. Zhang, and Z. Liang, "Application of BP Neural Network and Convolutional Neural Network (CNN) in Bearing Fault Diagnosis," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1043, no. 4, 2021, doi: 10.1088/1757-899X/1043/4/042026.

[19] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, "Plant leaf

disease detection using computer vision and machine learning algorithms," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 305–310, 2022, doi: 10.1016/j.gltp.2022.03.016.

[20] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network," *Procedia Comput. Sci.*, vol. 167, pp. 293–301, 2020, doi: 10.1016/j.procs.2020.03.225.

[21] H. I. Fitriasari and M. Rizkinia, "Improvement of Xception-ResNet50V2 Concatenation for COVID-19 Detection on Chest X-Ray Images," *2021 3rd East Indones. Conf. Comput. Inf. Technol.*, pp. 343–347, 2021, doi: 10.1109/EIConCIT50028.2021.9431916.

[22] D. Choi, C. J. Shallue, Z. Nado, J. Lee, C. J. Maddison, and G. E. Dahl, "On Empirical Comparisons of Optimizers for Deep Learning," no. 1, 2019, [Online]. Available: http://arxiv.org/abs/1910.05446

[23] Susanti, Mustakim, R. Novita, and I. Permana, "Application of Residual Network Architecture on Covid-19 Chest x-ray Classification," *Proceeding - 2022 Int. Symp. Inf. Technol. Digit. Innov. Technol. Innov. Dur. Pandemic, ISITDI 2022*, pp. 121–125, 2022, doi: 10.1109/ISITDI55734.2022.9944525.

[24] B. M. Abuhayi and A. A. Mossa, "Coffee disease classification using Convolutional Neural Network based on feature concatenation," *Informatics Med. Unlocked*, vol. 39, pp. 1–9, 2023, doi: 10.1016/j.imu.2023.101245.

[25] S. Zhong, J. Hu, X. Yu, and H. Zhang, "Molecular image-convolutional neural network (CNN) assisted QSAR models for predicting contaminant reactivity toward OH radicals: Transfer learning, data augmentation and model interpretation," *Chem. Eng. J.*, vol. 408, pp. 1–10, 2021, doi: 10.1016/j.cej.2020.127998.