

Classification of Cassava Leaf Diseases Using ResNet50 CNN Architecture Based on Digital Images

Maulana Malik^{1*}, Novan Wijaya²

^{1,2} Universitas Multi Data Palembang, Indonesia

¹maulanamalik_03@mhs.mdp.ac.id, ²novan.wijaya@mdp.ac.id



*Corresponding Author

Article History:

Submitted: 01-01-2026

Accepted: 04-01-2026

Published: 19-01-2026

Keywords:

Cassava; CNN; ResNet50; image classification; leaf disease.

Brilliance: Research of Artificial Intelligence is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

ABSTRACT

Cassava (*Manihot esculenta*) is an important agricultural commodity in Indonesia, but its productivity can decline due to leaf diseases such as Cassava Mosaic Disease (CMD), Cassava Green Mottle (CGM), Cassava Bacterial Blight (CBB), and Cassava Brown Streak Disease (CBSD). These four diseases exhibit overlapping visual symptoms such as chlorosis, spots, and leaf discoloration, making them difficult to distinguish manually. This study aims to create a digital-based cassava leaf image classification system using the Convolutional Neural Network (CNN) algorithm and ResNet50 architecture. The dataset used consists of 9,436 cassava leaf images taken from the TensorFlow platform and processed through resizing, normalization, selective augmentation, and the application of transfer learning. The experiment compared various optimizer configurations, learning rates, batch sizes, and balanced and unbalanced dataset scenarios. The evaluation was conducted using accuracy, precision, recall, and F1-score. The results show that the best performance was obtained on an unbalanced dataset using the Adam optimizer (learning rate 0.001; batch size 64) with an accuracy of 80.69% and an F1-score of 79.76%. Meanwhile, balancing the dataset actually reduced performance to an accuracy of 77.14% and an F1-score of 76.48%. Analysis of the loss curve and confusion matrix confirmed that the natural data distribution provided more stable generalization, although misclassification still occurred in classes with similar visual symptoms. These findings indicate that ResNet50 is effective for classifying cassava leaf diseases and has the potential to support early detection in digital agriculture practices.

INTRODUCTION

Cassava, commonly known in Indonesia as singkong (*Manihot esculenta* crantz), is known to have first grown in South America, specifically in what is now Brazil and Paraguay, during prehistoric times in Indonesia (Aisya & Prasetiadi, 2023). Cassava is one of the most important food crops in Indonesia. The tubers obtained from cassava plants are high in carbohydrates and can be cooked. Cassava leaves can also be used as vegetables. In addition, cassava can also be used as animal feed and industrial raw material (Arafat, Ichsan, & Pramoedya, 2025).

One of the causes of decline in production in terms of quantity and quality is disease. Cassava diseases can usually be detected through changes in parts of the plant, such as the leaves and stems (Nugraha, Rizal, & Pratiwi, 2022). Some of the main diseases that often infect cassava leaves include Cassava Mosaic Disease (CMD), Cassava Green Mite (CGM), Cassava Bacterial Canker (CBB), and Cassava Brown Streak Disease (CBSD). These four diseases show similar visual symptoms, such as spots, chlorosis, and damage to leaf tissue, which makes manual identification in the field difficult. The quality of cassava, whether good or bad, can affect the selling value of the cassava plant itself because the classification of diseases in cassava plants through laboratories requires a long time, is expensive, and laboratories are limited in availability (Arafat et al., 2025).

As artificial intelligence technology develops, deep learning-based approaches, particularly Convolutional Neural Networks (CNN), are becoming increasingly prevalent. CNNs can receive image inputs and determine what aspects or objects are present in the image, which machines can use to “learn” to recognize images and distinguish between them (Iswantoro & Handayani UN, 2022). CNNs utilize an artificial neural network architecture inspired by how the human visual system operates (Amalia, 2024).

In a study conducted by (Arafat et al., 2025), the MobileNet architecture applied in previous studies prioritized efficiency and compact model size, but had limitations in terms of depth and visual feature extraction capabilities. MobileNet does not have skip connections like ResNet50, making it less effective in recognizing complex visual patterns such as leaf spots that are similar in various diseases. This limitation is evident in the model's accuracy, which only reached 84%, without additional analysis such as precision and F1-score. Furthermore, research conducted by Nugraha et al. (2022), which applied the VGG16 architecture, only achieved a validation accuracy of 75%, with very low F1-scores in several classes such as CBB (0.32) and CGM (0.50). This indicates the limitations of VGG16 in identifying similar



visual symptoms on cassava leaves. In addition, the absence of skip connections in this architecture makes the training process more unstable. Therefore, this study utilizes ResNet50, a deeper architecture with a residual learning mechanism, to recognize complex visual patterns more accurately and consistently.

The ResNet50 architecture, known for its high performance in image classification and ability to handle vanishing gradient problems with its residual structure, was applied. Based on a study conducted by (Kulsum & Cherid, 2023) on the application of convolutional neural networks in plant classification using ResNet50, it can be concluded that the application of CNN with the ResNet50 model is very efficient because it is capable of classifying and predicting apple leaves consisting of two types of leaves, namely healthy apple leaves and rotten apple leaves, with an accuracy rate of 91%. Therefore, the application of the ResNet50 architecture for classification using images is a potential solution to improve efficiency and accuracy in detecting leaf spot disease on cassava leaves.

LITERATURE REVIEW

Several previous studies have conducted classification research on plant diseases using one of the most popular deep learning techniques for pattern recognition and image classification, namely convolutional neural networks (CNN). Because the model has thousands to millions of parameters, convolutional neural networks (CNN) are used to perform image-based classification (Hatur & Sabri, 2024). Convolutional Neural Network (CNN) is an architecture capable of identifying predictive indicators of an object in the form of images, text, voice recordings, and so on (Aisya & Prasetiadi, 2023).

The use of CNN architecture for classifying diseases in cassava leaves has been widely researched. This shows that the use of CNN architecture in classifying cassava leaf diseases is common because it is effective in performing classification. Some examples of cassava leaf research using CNN architecture are the study by (Nugraha et al., 2022) entitled "Classification of Diseases in Cassava Plants Using VGGNET Architecture Based on Deep Learning" and the study by (Lianardo, Syamsul, & Pratiwi, 2022) with the title "Classification of Leaf Disease Symptoms in Cassava Plants Based on Vision Using the CNN Method with Mobilenet Architecture", research by (Fahrezantara, Rizal, & Pratiwi, 2022) entitled "Utilization of Convolutional Neural Network in Cassava Plant Disease Classification Using Densenet Architecture", and research by (Arafat et al., 2025) with the title "Utilization of MOBILENETCNN Architecture to Diagnose Diseases in Cassava Leaves Through Digital Image Technology".

There has been no specific research in Indonesia that uses ResNet50 without mixing other CNN architectures to classify cassava leaf diseases. Residual Network50, commonly known as ResNet50, is a CNN architecture that has 50 layers (Fathur Rozi, Adiwijaya, & Swasono, 2023). Many studies show that, compared to other CNN architectures, ResNet50 has excellent performance. For example, in a study by Kulsum & Cherid (2023), it can be concluded that the use of CNNs using the ResNet50 model is very effective because CNNs can produce an accuracy value of 91% in the classification and prediction of apple leaf image data. Research conducted by (Fachri & Simbolon, 2024) obtained the highest accuracy results, reaching 100% accuracy by testing 0-41 epochs. Another example is in the research by (Suprihanto, Awaludin, Fadhil, & Zulfikor, 2022) The results of the research conducted on the ResNet-50 model in binary class and multiclass cases showed that it had the best performance in binary class cases with an accuracy of 92.68% and an f1-score value of 92.88%, while in multiclass cases, the best performance only achieved an accuracy of 88.98% and an f1-score value of 88.44%. Research conducted by (Maylianti, Ngurah, Wijayakusuma, Chandra, & Wiguna, 2025) shows that these two CNN architectures show the best results with an accuracy of 96%, each with a dataset division of 80% for training data and 20% for validation data. In terms of precision, the EfficientNetB0 architecture was superior with a value of 95.7%, while in terms of recall and f1 values, ResNet50 was superior with values of 94.7% and 93.3%. Another study using the ResNet50 architecture was conducted by (Made, Amanda, Gandhiadi, & Lanang, 2025) on two architectures, MobileNet-V2 and ResNet-50 V2, showing that MobileNet-V2 achieved an accuracy of 98.65%, while ResNet-50 V2 achieved 98.48%. It can be concluded that MobileNet -V2 is more resource-efficient, while ResNet50 V2 is more consistent in training deep networks. The last example is a study conducted by (Ramadhani et al., 2025) on a classification system for various diseases in mango leaves, showing that EfficientNetV2-S is more suitable for training with large datasets, while ResNet50 is ideal for smaller datasets.

METHOD

This study carried out several clearly defined steps, as illustrated in the flowchart in Figure 1. The first process began with a literature study on CNN with Resnet50 architecture for classifying cassava leaf diseases. This study carried out several clearly defined steps, as illustrated in the flowchart in Figure 1. The first process began with a literature study on CNN with Resnet50 architecture to classify cassava leaf diseases.

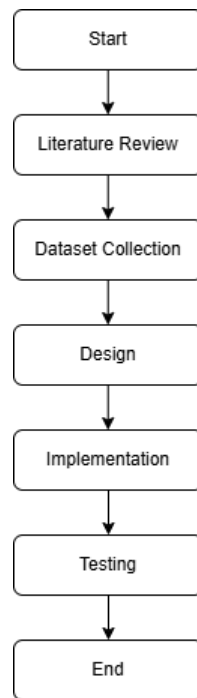


Fig 1. Methodology

The first stage is to begin and conduct a literature study. At this stage, various references related to the research topic are collected and analyzed, such as academic journals, books, and articles from leading publications. This study aims to understand the basic concepts and theories of image classification, Convolutional Neural Networks (CNN), ResNet50 architecture, and leaf spot diseases in cassava (Cassava Mosaic, Cassava Green Mite, Cassava Bacterial Canker, and Cassava Brown Streak Disease).

The next stage is the dataset collection process using a secondary dataset available from a public source in TensorFlow Datasets entitled “Cassava Disease Classification” which was introduced at the CVPR 2019 conference. This dataset consists of 9,436 annotated cassava leaf images divided into 5 classes (CBB, CBSD, CGM, CMD, Healthy). The labeling was done by experts from NaCRRI, making this dataset highly scientifically valid.

Table 1. Dataset class distribution

Number	Category	Count	Example Image
1	Cassava Bacterial Blight (CBB)	777	
2	Cassava Brown Streak Disease (CBSD)	2406	

3	Cassava Green Mottle (CGM)	1289
4	Cassava Mosaic Disease (CMD)	4431
5	Healty	527



The next step is classification design, where the ResNet50 architecture is applied using the transfer learning method, which uses the initial weights of the ResNet50 model that has been trained on a large general image dataset. The goal is for the model to have a basic understanding of common visual patterns, so that the training process can be more efficient and quickly adapt to the cassava leaf dataset. The images in this dataset vary in size and must be processed to be consistent with the model input. Therefore, all images are resized to 224 x 224 pixels. In addition, a new balanced dataset is created from the existing dataset using selective augmentation (flipping, rotation, and lighting variation) and the use of class weights in the model.

After implementing the previously designed system so that it can recognize and classify previously processed training data with a predetermined model, a program will be written to develop, train, and test the CNN ResNet50 model. This process will utilize the Python programming language, supported by the leading deep learning framework TensorFlow, to take advantage of libraries and functions that are efficient in deep learning calculations.

After the model has been successfully trained, the next step is to test and evaluate the model's performance using separate test data. After the testing phase, the test results are analyzed to determine the success rate of the applied method using a Confusion Matrix to calculate the precision, recall, and accuracy values, which can be seen in Equations (1), (2), and (3).

$$Precision = \frac{TP}{TP+FP} \times 100 \% \quad (1)$$

$$Recal = \frac{TP}{TP+FN} \times 100\% \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (3)$$

$$FI - Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (4)$$

RESULT

Table 2. Performance Comparison of ResNet50 Using Different Optimizers and Training Configurations

Optimizer	Learning rate	Batch Size	Dataset unbalanced				Dataset balanced			
			Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
RMSProp	0,001	16	80,32%	79,62%	80,32%	79,71%	69,55%	74,77%	69,55%	70,68%
		32	80,16%	79,44%	80,16%	79,23%	77,14%	77,53%	77,14%	76,48%
		64	80,32%	79,45%	80,32%	79,27%	73,00%	77,42%	73,00%	74,08%
	0,0001	16	77,45%	76,14%	77,45%	76,00%	73,58%	75,49%	73,58%	74,18%
		32	76,13%	74,76%	76,13%	74,56%	72,68%	75,13%	72,68%	73,39%
		64	74,01%	72,49%	74,01%	72,12%	70,45%	73,74%	70,45%	71,39%
Adam	0,001	16	79,79%	79,64%	79,79%	79,34%	76,18%	77,68%	76,18%	76,65%
		32	79,68%	79,34%	79,68%	78,56%	74,11%	77,00%	74,11%	74,86%
		64	80,69%	80,06%	80,69%	79,76%	75,92%	77,07%	75,92%	76,35%
Loss Curve	0,0001	16	77,51%	76,69%	77,51%	76,23%	73,63%	75,16%	73,63%	74,02%
		32	77,40%	76,38%	77,40%	76,19%	72,63%	74,44%	72,63%	73,01%
		64	75,28%	74,31%	75,28%	73,76%	70,34%	73,44%	70,34%	71,27%

As presented in Table 2, the highest performance on the unbalanced dataset was achieved using the Adam optimizer with a learning rate of 0.001 and a batch size of 64, resulting in an accuracy of 80.69% and an F1-score of 79.76%. In contrast, the best performance on the balanced dataset was obtained using the RMSProp optimizer with a learning rate of 0.001 and a batch size of 32, achieving an accuracy of 77.14% and an F1-score of 76.48%, indicating a slight performance decrease after dataset balancing.

DISCUSSION

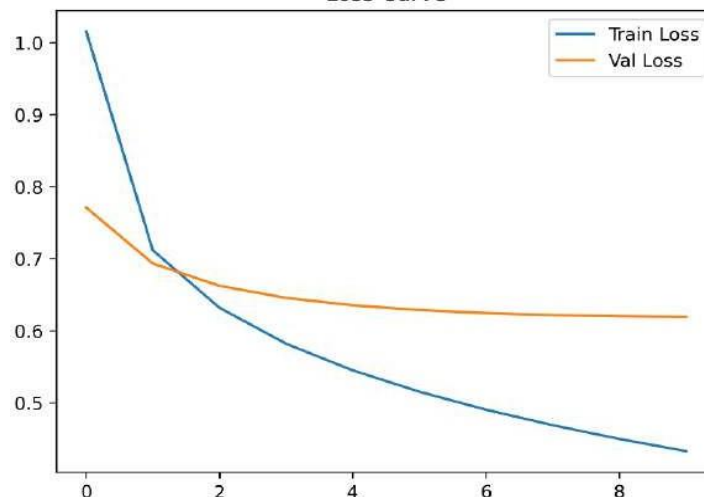


Fig. 2 Loss Curve of Best performance in Unbalanced Dataset

Figure 2 illustrates the training and validation loss curves of the ResNet50 model trained on the original (unbalanced) dataset. The training loss shows a consistent and steady decrease across epochs, indicating that the model successfully learns discriminative features from the training data. In contrast, the validation loss decreases rapidly during the early epochs and then gradually stabilizes, with a slight gap observed between training and validation loss toward the final epochs.

This behavior suggests that the model achieves convergence without severe overfitting. Although the gap between training and validation loss indicates a certain degree of generalization limitation, the validation loss does not increase

significantly, implying that the learned features remain stable and applicable to unseen data.

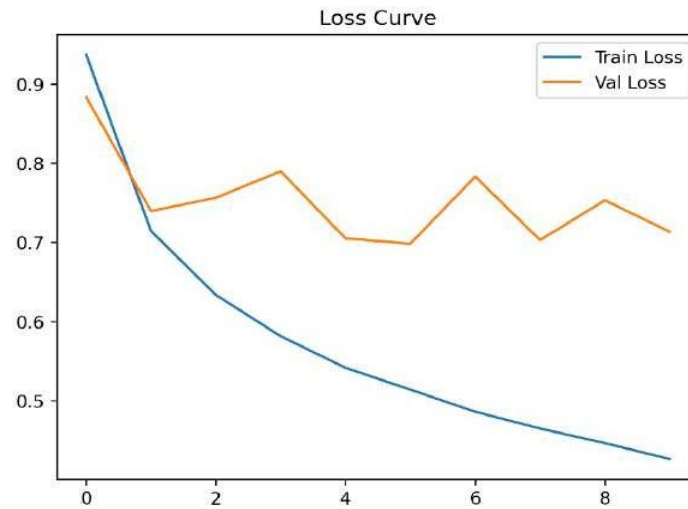


Fig. 3 Loss Curve of Best performance in Balanced Dataset

Figure 3 shows the training and validation loss curves of the ResNet50 model trained using the balanced dataset. The training loss decreases steadily across epochs, indicating that the model is able to learn the training data effectively after the balancing process. This consistent downward trend reflects stable optimization during the training phase.

In contrast, the validation loss exhibits noticeable fluctuations throughout the training process and does not follow a strictly decreasing pattern. After an initial decrease, the validation loss oscillates across subsequent epochs, suggesting instability in the model's generalization performance. This behavior indicates that, although the model fits the balanced training data well, it struggles to consistently generalize to the validation set.

The gap between training and validation loss becomes more pronounced compared to the original dataset, which suggests a higher tendency toward overfitting when using the balanced dataset. This phenomenon may be caused by the reduction of natural data diversity or the introduction of synthetic patterns during the balancing process, which can limit the model's ability to capture representative visual features of cassava leaf diseases.

When compared to the loss curve obtained from the original dataset, the balanced dataset results show reduced training stability and weaker generalization, as reflected by the fluctuating validation loss. This observation supports the quantitative findings, where the balanced dataset achieved lower accuracy and F1-score than the unbalanced dataset. It indicates that preserving the original data distribution allows the ResNet50 model to learn more robust and representative features, particularly for dominant disease classes.

Confusion Matrix

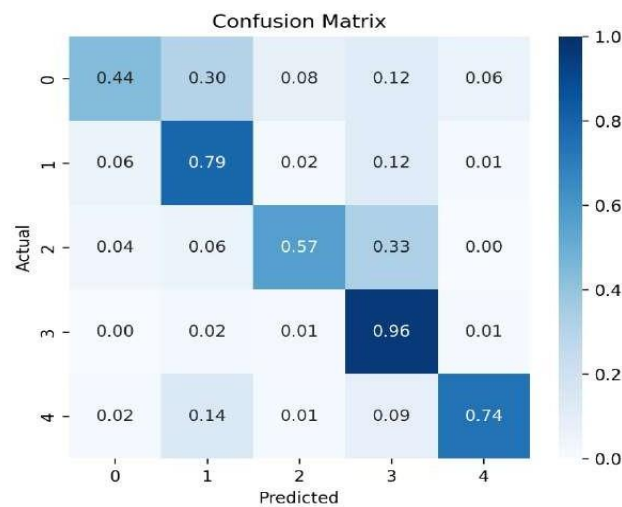


Fig. 4 Confusion Matrix of Best Performance in Unbalanced Dataset

Figure 4 presents the normalized confusion matrix for the ResNet50 model evaluated on the original (unbalanced) dataset. The diagonal elements indicate the proportion of correctly classified samples for each class. The model achieves the highest classification performance on class 3, with a correct prediction rate of 96%, demonstrating strong feature separability for this disease category. Class 1 and class 4 also exhibit relatively high recognition rates of

79% and 74%, respectively.

However, lower performance is observed for class 0 and class 2, which achieve correct classification rates of 44% and 57%. A noticeable misclassification occurs between class 0 and class 1, as well as between class 2 and class 3. This indicates that these classes share similar visual characteristics, such as leaf discoloration patterns and texture variations, making them more difficult to distinguish.

Despite these misclassifications, the confusion matrix confirms that the model performs well on dominant classes and maintains reasonable discrimination across all categories in the original dataset.

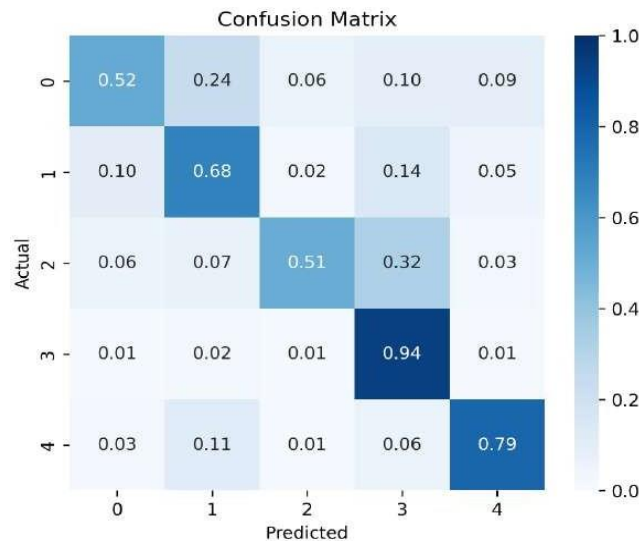


Fig. 5 Confusion Matrix of Best Performance in Balanced Dataset

Figure 5 illustrates the normalized confusion matrix of the ResNet50 model evaluated on the balanced dataset. The diagonal elements represent the proportion of correctly classified samples for each class, indicating the model's per-class recognition performance under balanced class distribution.

The model demonstrates its highest classification accuracy for class 3, achieving a correct prediction rate of approximately 94%, which indicates strong discriminative capability and clear feature separability for this category. Class 4 also shows robust performance with a recognition rate of 79%, while class 1 achieves a moderate accuracy of 68%, reflecting reasonable generalization when class imbalance is mitigated.

In contrast, class 0 and class 2 exhibit comparatively lower classification performance, with correct prediction rates of 52% and 51%, respectively. Significant confusion is observed between class 0 and class 1, as well as between class 2 and class 3. These misclassifications suggest that these class pairs share similar visual characteristics, such as overlapping leaf discoloration patterns, shape similarities, and subtle texture variations, which challenge the model's ability to distinguish between them even under balanced conditions.

Overall, the confusion matrix indicates that balancing the dataset improves fairness in class-wise evaluation and prevents dominance by majority classes. While the model maintains strong performance on visually distinctive categories, the remaining confusion highlights the need for enhanced feature extraction or complementary approaches, such as attention mechanisms or domain-specific feature augmentation, to further improve discrimination among visually similar disease classes.

CONCLUSION

Based on the experiments conducted, the implementation of the ResNet50 architecture for cassava leaf disease classification using digital images demonstrates satisfactory performance in distinguishing four target classes, namely CBB, CBSD, CGM, and CMD. The best performance was achieved using the unbalanced dataset with the Adam optimizer, learning rate of 0.001, and batch size of 64, resulting in an accuracy of 80.69% and an F1-score of 79.76%. Meanwhile, balancing the dataset led to a slight decrease in performance, with the highest accuracy of 77.14% and F1-score of 76.48%. These findings indicate that the natural data distribution provides richer and more representative visual features for the ResNet50 model.

The loss curve analysis further supports this result, showing that the model trained on the original dataset exhibits a more stable learning process without significant overfitting. In contrast, the balanced dataset presented notable fluctuations in validation loss, suggesting weaker generalization capability. The confusion matrix also shows strong recognition in dominant and visually distinctive classes, although several misclassifications still occur, particularly between classes with similar leaf symptom characteristics such as CBB–CMD and CGM–CBSD.

This study contributes to the development of automatic cassava disease detection systems using deep learning,



which can support digital agriculture applications by assisting farmers and agricultural practitioners in faster and more efficient disease identification. However, several limitations remain, including reliance on dataset quality, variations in lighting conditions, shooting angles, and the presence of visually similar disease symptoms that challenge the model.

For future work, it is recommended to explore more advanced or hybrid architectures, integrate attention mechanisms, apply more adaptive balancing techniques, and evaluate performance using real-field datasets to enhance robustness and reliability. Furthermore, developing a real-time implementation through mobile or IoT platforms has strong potential to expand the practical benefits of this system in modern agricultural practices.

REFERENCES

- Aisya, J. R., & Prasetiadi, A. (2023). Klasifikasi Penyakit Daun Kentang dengan Metode CNN dan RNN. *Jurnal Tekno Insentif*, 17(1), 1–10. <https://doi.org/10.36787/jti.v17i1.888>
- Amalia, D. F. (2024). Media Teknologi dan Informatika Pengolahan Citra Menggunakan Metode Convolutional Neural Network (CNN) Media Teknologi dan Informatika. 1, 1–8.
- Arafat, F. A., Ichsan, M. N., & Pramodya, M. F. (2025). Pemanfaatan Arsitektur MOBILENET-CNN Untuk Mendiagnosis Penyakit Pada Daun Singkong Melalui Teknologi Citra Digital. 4, 73–78.
- Fachri, H., & Simbolon, S. (2024). Classification bootstrap resnet50 for potato diseases datasets. *Jurnal Mantik*, 7(4), 2685–4236. Retrieved from <https://iocscience.org/ejournal/index.php/mantik/article/view/4587>
- Fahrezantara, A., Rizal, S., & Pratiwi, N. K. C. (2022). Pemanfaatan Convolutional Neural Network Dalam Klasifikasi Penyakit Tanaman Singkong Menggunakan Arsitektur Densenet. *EProceedings of Engineering*, 8(6), 3332–3338.
- Fathur Rozi, M. I., Adiwijaya, N. O., & Swasono, D. I. (2023). Identifikasi Kinerja Arsitektur Transfer Learning Vgg16, Resnet-50, Dan Inception-V3 Dalam Pengklasifikasian Citra Penyakit Daun Tomat. *Jurnal Riset Rekayasa Elektro*, 5(2), 145. <https://doi.org/10.30595/jrre.v5i2.18050>
- Hatur, Y., & Sabri, A. (2024). Perbandingan Arsitektur MobileNetV2 dan DenseNet121 untuk Klasifikasi. *Jurnal Ilmiah Komputasi*, 23(1), 67–74. <https://doi.org/10.32409/jikstik.23.1.3502>
- Iswantoro, D., & Handayani UN, D. (2022). Klasifikasi Penyakit Tanaman Jagung Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 900. <https://doi.org/10.33087/jiubj.v22i2.2065>
- Kulsum, U., & Cherid, A. (2023). Penerapan Convolutional Neural Network Pada Klasifikasi Tanaman Menggunakan ResNet50. *Simkom*, 8(2), 221–228. <https://doi.org/10.51717/simkom.v8i2.191>
- Lianardo, A., Syamsul, R., & Pratiwi, N. K. C. (2022). Klasifikasi Gejala Penyakit Daun Pada Tanaman Singkong Berbasis Vision Menggunakan Metode CNN Dengan Arsitektur Mobilenet. *E-Proceeding of Engineering*, 8(6), 3176–3179. Retrieved from <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18980>
- Made, D., Amanda, S., Gandhiadi, G. K., & Lanang, I. G. N. (2025). Kajian Metode Transfer Learning untuk Identifikasi Tumbuhan Herbal Berbasis Lontar. *Usada Taru Pramana*. 14(1), 77–89.
- Maylianti, N. P., Ngurah, I. G., Wijayakusuma, L., Chandra, I. P., & Wiguna, A. (2025). *Comparison of EfficientNet-B0 and ResNet-50 for Detecting Diseases in Cocoa Fruit*. 9(1), 115–120.
- Nugraha, A. E., Rizal, S., & Pratiwi, N. K. C. (2022). Klasifikasi Penyakit Pada Tanaman Singkong Menggunakan Arsitektur VGGNET Berbasis Deep Learning. *E-Proceeding of Engineering*, 8(6), 3240–3246.
- Ramadhani, K. F., Tarigan, M., Informatika, T., Unggul, U. E., Jeruk, K., Barat, K. J., ... Mildew, P. (2025). Implementasi Metode Convolutional Ceural Network untuk Klasifikasi Penyakit Daun Mangga menggunakan Arsitektur EfficientNetv2-s dan ResNet50. 9(3), 4135–4143.
- Suprihanto, S., Awaludin, I., Fadhil, M., & Zulfikor, M. A. Z. (2022). Analisis Kinerja ResNet-50 dalam Klasifikasi Penyakit pada Daun Kopi Robusta. *Jurnal Informatika*, 9(2), 116–122. <https://doi.org/10.31294/inf.v9i1.13049>