

Ensemble Learning-based Potato Leaf Disease Classification Using DenseNet201 and MobileNetV2

Burhan Ahmad^{1*}, Alamsyah²

^{1,2} Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia

DOI: <https://doi.org/10.52465/joiser.v4i1.597>

Received 10 July 2025; Accepted 26 January 2026; Available online 26 January 2026

Article Info

Keywords:

Image-based Plant Disease Classification; PlantVillage Dataset; Feature-level Ensemble; DenseNet201; MobileNetV2

Abstract

Early and late blight are major threats to potato crops and can cause significant losses for farmers. Early disease classification is essential for quick and appropriate treatment. This study proposes an ensemble learning approach by combining DenseNet201 and MobileNetV2 architectures to classify potato leaf diseases from digital images. The dataset used consists of 2,152 potato leaf images and is processed through normalization, augmentation, and image resizing stages. The ensemble model was trained with optimized parameters and evaluated using accuracy, precision, recall, and F1-score. The test results showed an accuracy of 99.56%, with precision, recall, and F1-score values of 99.56% each. Demonstrated improved performance compared to single CNN models on the evaluated dataset, and offers an accurate and efficient solution for disease detection in the agricultural sector.



This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

1. Introduction

The agricultural sector plays an important role in realizing global food security amidst the growing world population (Tabe-Ojong et al., 2024) (Costa et al., 2023). To respond to these dynamics, the agricultural sector is required to improve efficiency and productivity through various modern and innovative approaches (Gyamfi et al., 2024). Potato (*Solanum tuberosum*) is one of the strategic horticultural commodities that has high nutritional and economic value, especially in temperate to cold climates (MULUGETA et al., 2020) (Sholihati et al., 2020). In the context of food security, the success of potato cultivation determines the stability of food supply, given its large contribution to global consumption volumes (Tiwari et al., 2020). However, the productivity of potato plants is often hampered by foliar diseases such as early blight and late blight (DONG & ZHOU, 2022) (Radwan et al., 2024). Early blight is caused by *Alternaria solani*, while late blight is caused by *Phytophthora infestans*, both of which can cause heavy losses if not detected early (Meno et al., 2022) (Paluchowska et al., 2022). Manual disease detection is still common, but this method is subjective, slow and error-prone (Khan et al., 2025) (Arshaghi et al., 2023).

* Corresponding Author:

Burhan Ahmad,
Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang,
Semarang, Indonesia.
Email: bur3112003@students.unnes.ac.id

Along with the advancement of digital technology, artificial intelligence (AI), particularly artificial neural networks, has been increasingly applied for disease prediction (Alamsyah & Fadila, 2021). In particular, deep learning methods based on Convolutional Neural Networks (CNN) have begun to be used in plant disease classification based on leaf images (Harakannanavar et al., 2022) (Yao et al., 2023). Various CNN architectures such as VGG19, InceptionV3, and EfficientNetB0 have been used in previous studies (Tiwari et al., 2020) (Dutta et al., 2024) (Upadhyay et al., 2024). However, each architecture has limitations such as high computational requirements, overfitting, and difficulties in distinguishing disease symptoms with high visual similarity (Zubair et al., 2025) (Goyal et al., 2024) (Pai et al., 2025).

To address this issue, this study proposes an ensemble learning approach that combines two CNN architectures, namely DenseNet201 and MobileNetV2. DenseNet201 is known for its dense connections between layers, which support complex feature extraction (Charisma & Adhinata, 2023), while MobileNetV2 excels in computational efficiency, making it suitable for implementation on devices with limited resources (Chen et al., 2021). However, most existing studies rely on a single CNN architecture, which may limit generalization performance when disease symptoms exhibit high visual similarity. Therefore, this study proposes a feature-level ensemble approach for potato leaf diseases (early blight, late blight, and healthy) based on ensemble learning from digital images. The dataset used consists of 2,152 images of potato leaves with various visual conditions. It is hoped that this approach can enhance the accuracy and efficiency of plant disease classification systems in the agricultural sector.

2. Literature Review

Several previous studies have explored the use of machine learning and deep learning algorithms in the classification of potato leaf diseases. [23] used an image processing approach and the Random Forest algorithm on a dataset of images from PlantVillage, with the model achieving an accuracy of 97%. The process includes segmentation, feature extraction, and texture-based classification, but still relies on manual processing and is inefficient for large scales.

[6] applied a CNN model with a VGG-19 architecture using transfer learning to extract features from potato leaf images, which were then classified using logistic regression. The model achieved an accuracy of 97.80%. Although accurate, the VGG-19 architecture requires high computational power and is prone to overfitting.

Research by [24] took an image processing approach with preprocessing, K-means segmentation, texture feature extraction using Gray Level Co-occurrence Matrix (GLCM), and classification with Support Vector Machine (SVM). With the dataset evenly divided into three classes, this model achieved an accuracy of 95.99%, but classification for the late blight class tended to be less optimal due to the high visual similarity among diseases. This indicates that conventional machine learning-based methods still have limitations in distinguishing diseases with very similar visual symptoms.

[25] developed a more complex approach through image segmentation using a combination of graph cut algorithms, SLIC superpixels, and Otsu thresholding in the Lab color space, as well as classification using SVM, ANN, K-NN, and Random Forest. The SVM model achieved the highest accuracy of 97.4% for disease type classification. Although accurate, the segmentation approach used is quite complex and inefficient when applied to large-scale automatic classification systems in the field.

CNN-based deep learning approaches have also been utilized. [16] trained an InceptionV3 model on 2,152 images of potato leaves, achieving an accuracy of 98.60%. InceptionV3 is known for its efficiency, yet it still has limitations in distinguishing similar visual symptoms. Similarly, [17] used EfficientNetB0 on the same dataset and achieved an accuracy of 99.05%, demonstrating high performance but limited to a single CNN architecture.

Based on the limitations of previous research approaches, such as the need for high computational power, overfitting, and difficulty in distinguishing diseases with high visual similarity, this study proposes an ensemble learning method combining DenseNet201 and MobileNetV2. This combination aims to integrate the advantages of deep visual features and computational efficiency, to significantly improve the accuracy and generalization of potato leaf disease classification.

3. Method

Research method is an important component in conducting a study. In this research, the method used is an ensemble learning approach by combining two CNN architectures, namely DenseNet201 and MobileNetV2. The workflow of this method is presented in the form of a flow chart shown in Figure 1.

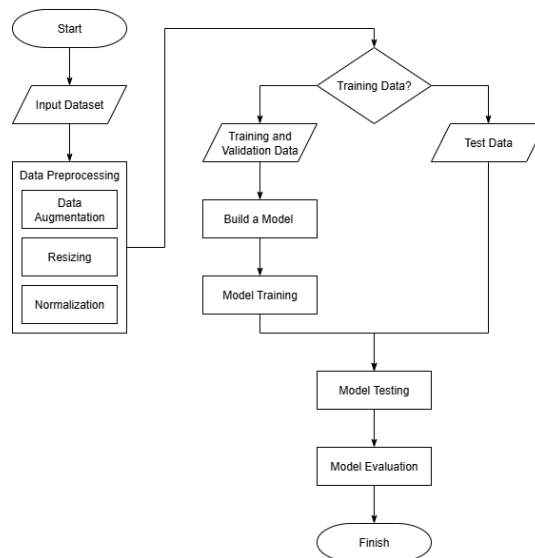


Figure 1. Research Procedure Flowchart

a. Input Data

The first step of this research begins with the process of inputting the dataset that will be used in training the classification model. The dataset is obtained from PlantVillage, which is a public dataset source with a subset of potato leaf images consisting of three classes. In total, there are 1,000 images for each of the early blight and late blight classes, and 152 images for the healthy leaf class.

b. Data Preprocessing

Data preprocessing is an important step before the model training process, which will process the data because the data collected is usually incomplete, redundant, and inconsistent (Alamsyah & Fadila, 2021). The data preprocessing stage is carried out to prepare the dataset to meet the input specifications of the CNN model and to improve training performance. The first step is to conduct an initial data exploration (EDA) to understand the distribution of classes and image characteristics, including size, lighting, and the number of images per class. Next, image augmentation is performed specifically on the imbalanced class, which is the healthy class. Augmentation techniques used include random rotation, zooming, brightness adjustment, and horizontal/vertical flipping (Prasetyo et al., 2023). The goal of this augmentation is to enhance the diversity of visual data, increase the number of samples, and minimize the risk of overfitting during the model training process. After that, all images are resized to 224×224 pixels to meet the specifications of the DenseNet201 and MobileNetV2 architectures. The final stage of preprocessing is normalizing the RGB pixels from a scale of 0–255 to a range of 0–1. This normalization aims to stabilize training and speed up convergence.

c. Dataset Splitting

The dataset is divided into three parts: training data, validation, and testing. The division is done with a ratio of 70% training data with 2,100 images, 15% for validation data, and 15% for testing data, 4 each with 450 images. The purpose of this division is to ensure that the developed model can generalize to new data that it has not seen before.

d. Model Method

The classification model in this research is designed using an ensemble learning approach based on feature-level fusion, specifically by combining two pre-trained Convolutional Neural Network (CNN) architectures, namely DenseNet201 and MobileNetV2. Fine-tuning was not performed to reduce computational complexity and to evaluate the effectiveness of feature-level fusion using pretrained representations. The purpose of utilizing these two backbones is to combine the strengths of each architecture in extracting important visual features from potato leaf images, to improve classification accuracy and strengthen model generalization.

Images with a size of 224×224 pixels with three color channels (RGB) are input into the model through the input layer, which is then processed in parallel by DenseNet201 and MobileNetV2. Each

architecture produces feature tensors with different dimensions based on the internal characteristics of each model. To efficiently combine features from both networks, dimensionality reduction is performed using a 2D Global Average Pooling layer, aimed at summarizing the spatial features from the CNN output into a one-dimensional vector without losing important information.

The pooling results from both CNN pathways are then combined using a concatenation layer to form a more informative combined feature representation. After merging, the data passes through a dropout layer with a ratio of 0.4 as a regularization mechanism to prevent overfitting during training. Subsequently, this combined vector is connected to a dense (fully connected) layer with three output neurons, corresponding to the number of target classes: early blight, late blight, and healthy. The Softmax activation function is used in this layer to produce a probability distribution as the final classification result. The model architecture used can be seen in Figure 2.

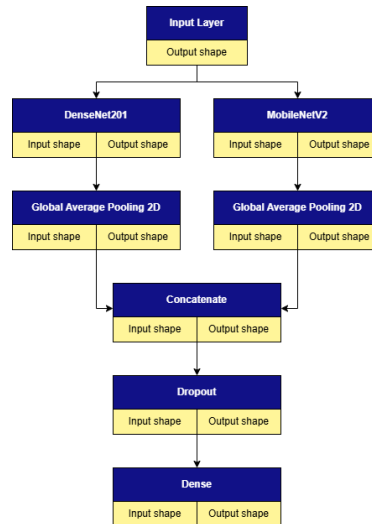


Figure 2. Model Architecture

The image depicts the data processing flow from the input layer to producing the final prediction. The parallel pathways of DenseNet201 and MobileNetV2 allow the model to capture features from different architectural perspectives, thus enriching the information extraction process. The use of Global Average Pooling, vector merging, dropout, and dense layers creates a compact yet powerful design. With this architecture, the model is capable of accurately and efficiently classifying potato leaf diseases.

e. Model Training

The training process was carried out to optimize the model's performance in classifying potato leaf diseases. The training included real-time data augmentation, forward pass, loss calculation, backpropagation, and weight updates using the Adam optimizer. The Adam optimizer was used with a learning rate of 0.0001. Early stopping was applied to prevent overfitting. The model was trained for 20 epochs with a batch size of 32, using the Softmax activation function and sparse categorical crossentropy as the loss function. During training, the model's performance was monitored through validation data to prevent overfitting. The training results in a weight file that is ready to be used for testing and implementing the classification system.

f. Model Testing

Model testing. The testing is conducted using test data that was not used during training or validation, in order to measure the model's generalization ability. The process is carried out to produce class predictions.

g. Model Evaluation

To evaluate the model performance, one can look at the evaluation metrics (Alamsyah et al., 2023). Examples of calculation of evaluation metrics include accuracy, precision, recall, and F1 score on the test set as shown in equations (1), (2), (3), and (4).

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

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (3)$$

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

4. Results and Discussion

Results

a. Model Training Results

The model training process was carried out for 20 epochs using training and validation data that had gone through the normalization stage. The model was compiled using Adam's optimizer with the main metric being accuracy. During training, the model performed very well from the first epoch, with a training accuracy of 86.32% and a validation accuracy of 97.54%. The accuracy value continued to increase at each epoch, while the loss decreased consistently. These results show that the model is able to learn visual patterns effectively without overfitting. A graph of the model's training performance is shown in Figure 3.

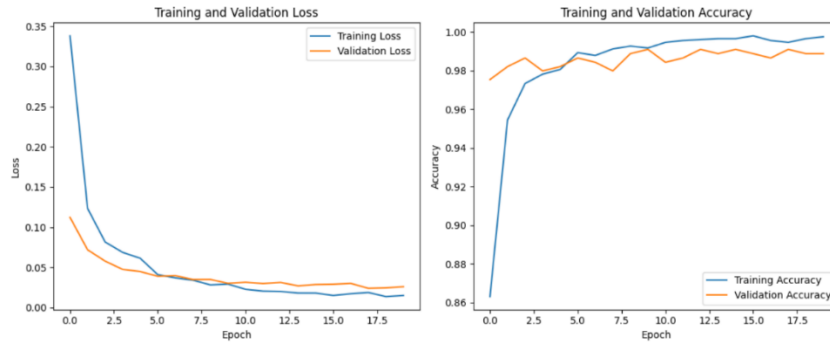


Figure 3. Model Performance During Training

b. Model Testing Results

Tests were conducted on test datasets that had not been used in the training stage, to evaluate the generalization ability of the ensemble model. The ensemble model based on DenseNet201 and MobileNetV2 achieved an accuracy of 99.56% on the test data, indicating a very high classification performance. These results prove the effectiveness of the feature-level ensemble approach in improving accuracy, while confirming the model's ability to generalize new visual patterns with excellent precision.

c. Model Evaluation Results

Evaluation of the model's performance on the test data was carried out by constructing a confusion matrix to measure the accuracy of classification in each class: early blight, healthy, and late blight. The visualization of the confusion matrix is shown in Figure 4.

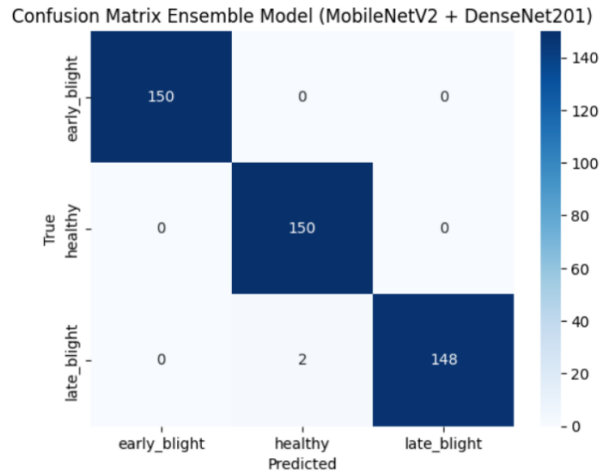


Figure 4. Confusion Matrix

The results show that the model successfully classified all early blight and healthy images correctly as many as 150 images each. For the late blight class, there were 2 images that were misclassified as healthy, while 148 others were correctly classified. This result shows a very low misclassification rate and good generalization performance. Next, the performance of the model was evaluated using classification metrics, namely precision, recall, F1-score, and accuracy. This evaluation was conducted using the classification report. The evaluation results are shown in Table 1.

Table 1. Classification Report Result

	Precision	Recall	F1-Score	Support
early_blight	1.0000	1.0000	1.0000	150
healthy	0.9868	1.0000	0.9934	150
late_blight	1.0000	0.9867	0.9933	150
accuracy			0.9956	450
macro avg	0.9956	0.9956	0.9956	450
weighted avg	0.9956	0.9956	0.9956	450

Discussion

In this study, the proposed method using a combination of two Convolutional Neural Network (CNN) architectures, namely DenseNet201 and MobileNetV2, is compared with a number of previous research models in the field of potato leaf disease classification. The purpose of this comparison is to evaluate the performance improvement offered by the ensemble approach used. Details of the model comparison are presented in Table 2.

Table 2. The Comparison with Previous Model

Author	Year	Model	Result
(Iqbal & Talukder, 2020)	2020	RF	Accuracy 97%
(Tiwari et al., 2020)	2020	CNN with VGG-19 Architecture	Accuracy 97.8%
(Singh & Kaur, 2021)	2021	SVM	Accuracy 95.99%
(Hou et al., 2021)	2021	KNN, SVM, ANN, and RF	The best accuracy of SVM is 97.40%.
(Dutta et al., 2024)	2024	CNN with InceptionV3 Architecture	Accuracy 98.6%
(Upadhyay et al., 2024)	2024	CNN with EfficientNetB0 Architecture	Accuracy 99.05%
Proposed Method	2025	CNN With Architecture Densenet201 and Mobilenetv2 based Ensemble Learning	Accuracy 99.56%

The proposed method shows the best performance compared to previous studies by combining the strengths of DenseNet201 in maintaining information flow and MobileNetV2 which is computationally

efficient. Through the Feature-Level Ensemble approach, the model was able to achieve 99.56% accuracy, as well as near-perfect precision, recall, and F1-score values. The combination of these two pretrained models significantly improved the classification ability of the three potato leaf disease classes (early blight, late blight, and healthy). Thus, the proposed model provides a more optimized and accurate solution and has high potential to be applied in artificial intelligence-based image classification systems in precision agriculture. Despite the promising results, this study has limitations. The dataset was collected under controlled conditions, which may not fully represent real-world agricultural environments.

5. Conclusion

Potato leaf diseases are one of the main factors affecting the productivity and quality of agricultural produce. Manual methods of disease detection are still widely used, but have disadvantages such as dependence on human expertise and potential identification errors. Therefore, this study developed an image-based automatic classification model using an ensemble learning approach that combines two Convolutional Neural Network (CNN) architectures, namely DenseNet201 and MobileNetV2. The results showed that the Feature-Level Ensemble approach that was applied succeeded in significantly improving the performance of the model. The model was able to achieve an accuracy of 99.56%, with precision, recall, and F1-score values above 99% for each class, namely early blight, late blight, and healthy. The model also showed high stability during training without overfitting, and was able to generalize well to the test data. With a high level of accuracy and efficiency, this model has the potential to be applied in real-time plant disease classification systems, both in the form of web applications and artificial intelligence-based monitoring systems in agricultural environments. However, further testing of images from varying field conditions is required to ensure the robustness of the model in real-world scenarios.

References

- Alamsyah, A., & Fadila, T. (2021). Increased accuracy of prediction hepatitis disease using the application of principal component analysis on a support vector machine. *Journal of Physics: Conference Series*, 1968(1). <https://doi.org/10.1088/1742-6596/1968/1/012016>
- Alamsyah, Prasetyo, B., Hakim, M. F. Al, & Pradana, F. D. (2023). The improvement of COVID-19 prediction accuracy using optimal parameters in recurrent neural network model. *AIP Conference Proceedings*, 2614(1), 40019. <https://doi.org/10.1063/5.0125767>
- Arshaghi, A., Ashourian, M., & Ghabeli, L. (2023). Potato Diseases Detection and Classification Using Deep Learning Methods. *Multimedia Tools and Applications*, 82(4), 5725–5742. <https://doi.org/10.1007/s11042-022-13390-1>
- Charisma, R. A., & Adhinata, F. D. (2023). Transfer Learning With Densenet201 Architecture Model For Potato Leaf Disease Classification. *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, 738–743. <https://doi.org/10.1109/ICCoSITE57641.2023.10127772>
- Chen, J., Zhang, D., Suzauddola, M., & Zeb, A. (2021). Identifying Crop Diseases Using Attention Embedded MobileNet-V2 Model. *Applied Soft Computing*, 113, 107901. <https://doi.org/https://doi.org/10.1016/j.asoc.2021.107901>
- Costa, F., Frecassetti, S., Rossini, M., & Portioli-Staudacher, A. (2023). Industry 4.0 Digital Technologies Enhancing Sustainability: Applications and Barriers From the Agricultural Industry in an Emerging Economy. *Journal of Cleaner Production*, 408(December 2022), 137208. <https://doi.org/10.1016/j.jclepro.2023.137208>
- DONG, S. meng, & ZHOU, S. qun. (2022). Potato Late Blight Caused by Phytophthora Infestans: From Molecular Interactions to Integrated Management Strategies. *Journal of Integrative Agriculture*, 21(12), 3456–3466. <https://doi.org/10.1016/j.jia.2022.08.060>
- Dutta, M., Gupta, D., Gulzar, Y., Mir, M. S., Onn, C. W., & Soomro, A. B. (2024). Leveraging Inception V3 for Precise Early and Late Blight Disease Classification in Potato Crops. *Traitement Du Signal*, 41(2), 705–715. <https://doi.org/10.18280/ts.410213>
- Goyal, P., Sharda, R., Saini, M., & Siag, M. (2024). A Deep Learning Approach for Early Detection of Drought Stress in Maize using Proximal Scale Digital Images. *Neural Computing and Applications*, 36(4), 1899–1913. <https://doi.org/10.1007/s00521-023-09219-z>
- Gyamfi, E. K., ElSayed, Z., Kropczynski, J., Yakubu, M. A., & Elsayed, N. (2024). Agricultural 4.0 Leveraging on Technological Solutions: Study for Smart Farming Sector. *2024 IEEE 3rd International*

- Conference on Computing and Machine Intelligence (ICMI)*, 1–9. <https://doi.org/10.1109/ICMI60790.2024.10585910>
- Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant Leaf Disease Detection Using Computer Vision and Machine Learning Algorithms. *Global Transitions Proceedings*, 3(1), 305–310. <https://doi.org/10.1016/j.gltp.2022.03.016>
- Hou, C., Zhuang, J., Tang, Y., He, Y., Miao, A., Huang, H., & Luo, S. (2021). Recognition of Early Blight and Late Blight Diseases on Potato Leaves Based on Graph Cut Segmentation. *Journal of Agriculture and Food Research*, 5, 100154. <https://doi.org/10.1016/j.jafr.2021.100154>
- Iqbal, M. A., & Talukder, K. H. (2020). Detection of Potato Disease Using Image Segmentation and Machine Learning. *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, 43–47. <https://doi.org/10.1109/WiSPNET48689.2020.9198563>
- Khan, S. U., Alsuhaibani, A., Alabduljabbar, A., Almarshad, F., Altherwy, Y. N., & Akram, T. (2025). A Review on Automated Plant Disease Detection: Motivation, Limitations, Challenges, and Recent Advancements for Future Research. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 37, Number 3). Springer International Publishing. <https://doi.org/10.1007/s44443-025-00040-3>
- Meno, L., Abuley, I. K., Escuredo, O., & Seijo, M. C. (2022). Suitability of Early Blight Forecasting Systems for Detecting First Symptoms in Potato Crops of NW Spain. *Agronomy*, 12(7), 8–12. <https://doi.org/10.3390/agronomy12071611>
- MULUGETA, T., MUHINYUZA, J. B., GOUWS-MEYER, R., MATSAUNYANE, L., ANDREASSON, E., & ALEXANDERSSON, E. (2020). Botanicals and Plant Strengtheners for Potato and Tomato Cultivation in Africa. *Journal of Integrative Agriculture*, 19(2), 406–427. [https://doi.org/10.1016/S2095-3119\(19\)62703-6](https://doi.org/10.1016/S2095-3119(19)62703-6)
- Pai, P., Amutha, S., Basthikodi, M., Ahamed Shafeeq, B. M., Chaitra, K. M., & Gurpur, A. P. (2025). A Twin CNN-based Framework for Optimized Rice Leaf Disease Classification With Feature Fusion. *Journal of Big Data*, 12(1). <https://doi.org/10.1186/s40537-025-01148-z>
- Paluchowska, P., Śliwka, J., & Yin, Z. (2022). Late Blight Resistance Genes in Potato Breeding. *Planta*, 255(6), 1–20. <https://doi.org/10.1007/s00425-022-03910-6>
- Prasetyo, B., Alamsyah, Hakim, M. F. Al, Jumanto, & Adi, M. H. (2023). Differential augmentation data for vehicle classification using convolutional neural network. *AIP Conference Proceedings*, 2614(1), 40001. <https://doi.org/10.1063/5.0126720>
- Radwan, M., Alhussan, A. A., Ibrahim, A., & Tawfeek, S. M. (2024). Potato Leaf Disease Classification Using Optimized Machine Learning Models and Feature Selection Techniques. *Potato Research*, (0123456789). <https://doi.org/10.1007/s11540-024-09763-8>
- Sholihati, R. A., Sulistijono, I. A., Risnumawan, A., & Kusumawati, E. (2020). Potato Leaf Disease Classification Using Deep Learning Approach. *IES 2020 - International Electronics Symposium: The Role of Autonomous and Intelligent Systems for Human Life and Comfort*, 392–397. <https://doi.org/10.1109/IES50839.2020.9231784>
- Singh, A., & Kaur, H. (2021). Potato Plant Leaves Disease Detection and Classification Using Machine Learning Methodologies. *IOP Conference Series: Materials Science and Engineering*, 1022(1). <https://doi.org/10.1088/1757-899X/1022/1/012121>
- Tabe-Ojong, M. P. Jr., Nana, I., Zimmermann, A., & Jafari, Y. (2024). Trends and Evolution of Global Value Chains in Food and Agriculture: Implications for Food Security and Nutrition. *Food Policy*, 127, 102679. <https://doi.org/https://doi.org/10.1016/j.foodpol.2024.102679>
- Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S., & Bhardwaj, S. (2020). Potato Leaf Diseases Detection Using Deep Learning. *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 461–466. <https://doi.org/10.1109/ICICCS48265.2020.9121067>
- Upadhyay, S. K., Jain, J., & Prasad, R. (2024). Early Blight and Late Blight Disease Detection in Potato Using Efficientnetb0. *International Journal of Experimental Research and Review*, 38, 15–25. <https://doi.org/10.52756/ijerr.2024.v38.002>
- Yao, J., Tran, S. N., Sawyer, S., & Garg, S. (2023). Machine Learning for Leaf Disease Classification: Data, Techniques and Applications. In *Artificial Intelligence Review* (Vol. 56, Number s3). Springer Netherlands. <https://doi.org/10.1007/s10462-023-10610-4>
- Zubair, F., Saleh, M., Akbari, Y., & Al Maadeed, S. (2025). A Robust Ensemble Model for Plant Disease Detection Using Deep Learning Architectures. *AgriEngineering*, 7(5). <https://doi.org/10.3390/agriengineering7050159>