Type of Contribution:
Research Paper
Review Paper
Case Study

INTRO: JURNAL INFORMATIKA DAN TEKNIK ELEKTRO
DOI: 10.51747/intro.v4i2.425
ISSN 3025-602X

This article contributes to:

SUSTAINABLE DEVELOPMENT GOALS
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

Automatic Identification of Tomato Leaf Conditions Based on OpenCV and Convolutional Neural Networks

Chaiden Richardo Foanto[1], **Leonard Widjaja**[1], **Alicia Juanita Lisal**[1], **Citra Suardi**[1*]

[1] Department of Informatics (Makassar City Campus), Ciputra University Makassar, 90224, **Indonesia**

*citra.suardi@ciputra.ac.id

Abstract

Tomato leaf diseases are a major cause of yield loss, particularly in rural areas with limited access to agricultural experts and diagnostic facilities. This study proposes an artificial intelligence–based system for identifying tomato leaf conditions using a Convolutional Neural Network (CNN) integrated with OpenCV. The system classifies tomato leaves into nine categories, including one healthy class and eight disease classes, based on digital images. The dataset was divided into training, validation, and testing sets, with pre-processing steps including resizing, normalization, and data augmentation. The CNN model was trained using the Adam optimizer and categorical cross-entropy loss. Experimental results show that the model achieved approximately 90% accuracy, with average precision, recall, and F1-score values above 0.88, indicating strong classification performance and good generalization ability. The OpenCV-based implementation enables real-time detection via a camera with an average prediction time of less than one second per image. These findings demonstrate that integrating CNN with OpenCV provides a practical and efficient solution for early tomato leaf disease detection and supports decision-making in technology-driven agriculture.

Keywords: CNN; Computer Vision; Deep Learning; Image Classification; OpenCV; Tomato Leaf Disease

Article Info
Submitted: 2025-07-03
Revised: 2025-10-15
Accepted: 2025-12-25

Publisher
Universitas Panca Marga

1. Introduction

Horticultural agriculture, particularly tomato cultivation, plays an important role in supporting rural economies in Indonesia, including in areas such as Sengka Village. However, tomato productivity is often threatened by rapidly spreading leaf diseases that are difficult to identify at an early stage by farmers. Symptoms such as brown spots, chlorosis, and necrosis are usually noticed only after infections become severe, leading to delayed control measures and significant yield losses [1].

Manual identification of plant diseases requires substantial technical knowledge of pathogen-specific visual symptoms, while many farmers in remote areas have limited access to agricultural extension services or laboratory diagnostic facilities. Laboratory-based methods such as PCR and microbial culture are impractical for daily use due to their high cost, time requirements, and infrastructure needs. As a result, there is a strong demand for fast, affordable, and easy-to-use disease detection tools for field conditions.

Recent advances in artificial intelligence, particularly in computer vision, offer a promising alternative. Convolutional Neural Networks (CNNs) have proven highly effective in plant image classification tasks, including the recognition of tomato leaf diseases based on visual patterns such as color, texture, and lesion shape [2]. CNNs automatically learn discriminative features from large image datasets, eliminating the need for manual feature extraction and making them suitable for real-world agricultural applications.

Integrating CNN models with image-processing libraries such as OpenCV enables the development of real-time disease detection systems using standard cameras, including low-cost mobile devices [3]. This approach supports precision agriculture by allowing farmers to obtain instant diagnoses simply by capturing images of plant leaves. Therefore, this study aims to develop a CNN-based tomato leaf condition identification system integrated with OpenCV to classify leaves into nine categories (one healthy class and eight disease classes). The proposed system is designed to be robust under varying lighting and background conditions commonly found in real farming environments and is expected to serve as a practical digital tool to support decision-making and improve tomato production resilience.

2. Methods

This study employs a deep learning approach using a Convolutional Neural Network (CNN) to classify tomato leaf diseases based on digital images. CNN was selected due to its strong capability in automatically learning visual patterns and features from images, making it widely used in computer vision–based image classification tasks [4, [5]. The methodology integrates two main stages: (1) CNN model training and (2) real-time disease detection using a computer vision system based on OpenCV.

In the training stage, a dataset of tomato leaf images was grouped into several classes according to disease categories. The dataset was divided into training, validation, and testing sets to ensure objective learning and evaluation [5]. Prior to training, all images were pre-processed through resizing and pixel-value normalization to match the CNN input requirements. Data augmentation techniques were applied to the training set to increase image variability and improve the model’s robustness under different visual conditions [6].

The CNN architecture consists of multiple convolutional (Conv2D) layers for feature extraction, followed by max-pooling layers to reduce dimensionality and computational complexity. Extracted features are flattened and passed to fully connected (dense) layers for classification. A softmax activation function is used in the output layer to generate class probabilities for each leaf condition [7]. The model was trained using the Adam optimizer with categorical cross-entropy loss for multi-class classification. Validation data were used to monitor performance and reduce overfitting. After training, the model was saved in HDF5 (.h5) format for deployment without retraining [8].

In the implementation stage, the trained CNN model was integrated into a real-time detection system using OpenCV and Python. A local camera provides live image input, which is pre-processed using the same resizing and normalization steps as in training. Each frame is then fed into the CNN model to generate disease predictions [9]. The system displays the predicted class label and confidence score directly on the camera interface, along with brief explanatory information and handling suggestions. This integration of CNN and OpenCV enables a practical computer vision–based decision support system that can be directly applied in real agricultural environments [10].

3.Results and Discussion

The training results of the Convolutional Neural Network (CNN) model indicate that the model gradually learned the patterns of plant leaf diseases as the number of epochs increased. This can be observed from the training and validation accuracy curves, which show a relatively consistent upward trend, suggesting that the model successfully adjusted its weights to capture the visual characteristics of leaf images. The small gap between training and validation accuracy further indicates that the model has good generalization ability and is not strongly affected by overfitting.

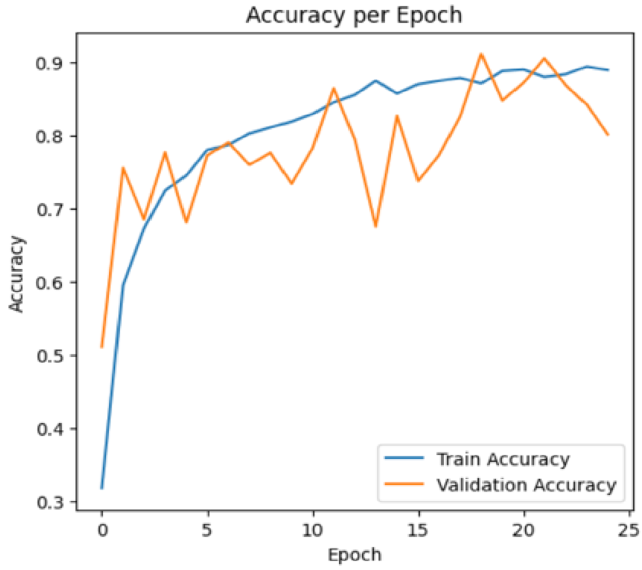

Figure 1. Graph of CNN Model Accuracy on Training and Validation Data Against the Number of Epochs

In addition to accuracy, the loss-per-epoch curves provide important insight into the stability of the model's learning process. The decreasing loss values for both the training and validation sets indicate that the model's prediction errors were progressively reduced as training proceeded. The relatively balanced loss patterns between training and validation further suggest that the model does not suffer from significant overfitting.

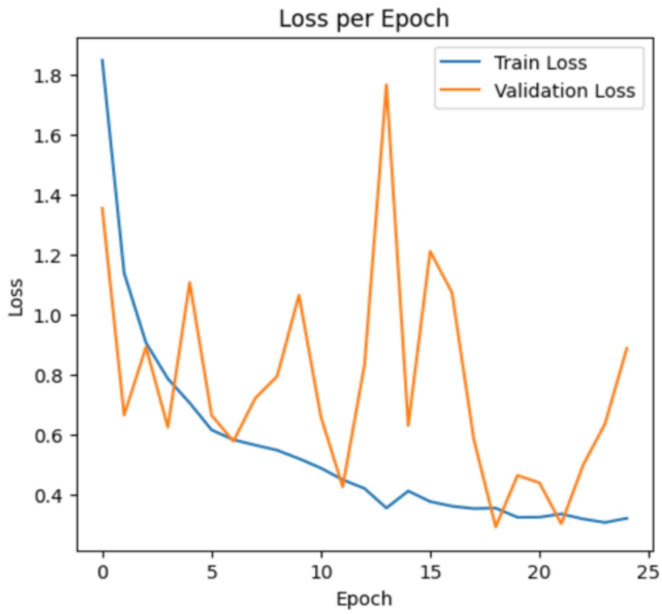

Figure 2. Graph of CNN Model Loss Values on Training and Validation Data Against the Number of Epochs

The model's performance was evaluated on the test dataset to assess its ability to classify leaf images that were not seen during training. A confusion matrix was used to present the detailed distribution of the model's predictions for each disease class. Through the confusion matrix, it is possible to identify which classes are most accurately classified and which ones still exhibit misclassification errors.

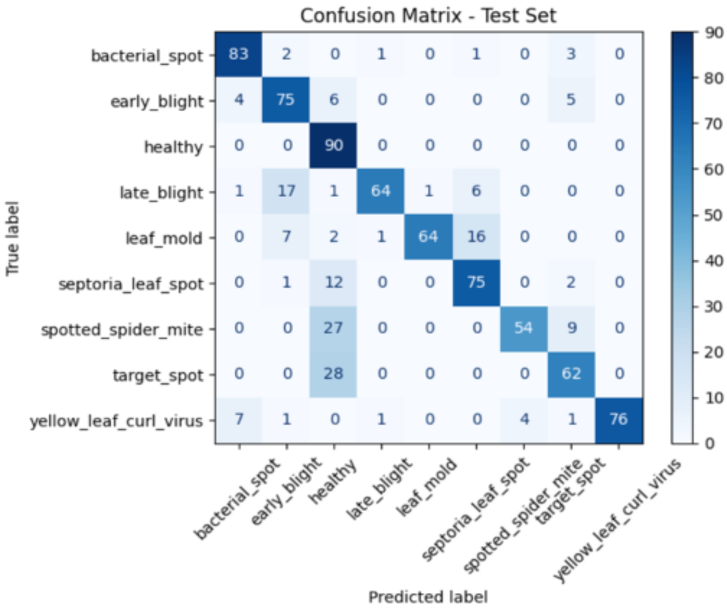

Figure 3. Confusion Matrix of CNN Model Evaluation Results Using Test Data

In addition to the confusion matrix, a classification report was used to evaluate the model's performance based on precision, recall, and F1-score for each leaf disease class. Precision reflects the accuracy of the model's predictions, while recall indicates the model's ability to correctly identify all instances of a given class. The F1-score represents the harmonic mean of precision and recall, providing a more comprehensive measure of the model's overall classification performance.

```
Classification Report:

                        precision    recall  f1-score   support

        bacterial_spot       0.87      0.92      0.90        90
          early_blight       0.73      0.83      0.78        90
               healthy       0.54      1.00      0.70        90
           late_blight       0.96      0.71      0.82        90
             leaf_mold       0.98      0.71      0.83        90
    septoria_leaf_spot       0.77      0.83      0.80        90
   spotted_spider_mite       0.93      0.60      0.73        90
           target_spot       0.76      0.69      0.72        90
yellow_leaf_curl_virus       1.00      0.84      0.92        90

              accuracy                           0.79       810
             macro avg       0.84      0.79      0.80       810
          weighted avg       0.84      0.79      0.80       810
```

Figure 4. Classification Report of CNN Model Based on Test Results on Test Data

In the computer vision implementation stage, the trained CNN model was successfully integrated with OpenCV to perform real-time leaf disease classification using a local camera. The system is able to capture leaf images directly, apply preprocessing, and generate predictions within a relatively short time. The prediction results are displayed on the camera interface in the form of the disease name and confidence level, allowing users to understand the classification output intuitively.

Figure 5. UI of OpenCV-based Computer Vision System for Real-Time Tomato Leaf Disease Classification

The system implementation results indicate that the proposed deep learning approach is not only effective in experimental settings but also practically applicable in real-world conditions. Although system performance can be influenced by external factors such as lighting conditions and image capture angles, the model is still able to produce relevant and reliable predictions. This demonstrates that CNN has strong potential to be further developed as a decision support system in artificial intelligence–based agriculture. This discussion reinforces the relevance of using CNN for image-based plant leaf disease classification [4].

4. Conclusion

This study successfully developed a tomato leaf condition identification system based on a Convolutional Neural Network (CNN) integrated with OpenCV for real-time disease classification. The training and testing results demonstrate that the CNN model effectively learns visual patterns of tomato leaves with good generalization ability, as reflected by stable accuracy trends, decreasing loss values, and satisfactory performance on the test dataset based on the confusion matrix and classification report. The computer vision implementation confirms that the model can be deployed using a standard camera with relatively fast prediction time and outputs that are intuitive for users. Therefore, the proposed system has strong potential as a practical, efficient, and AI-based decision support tool to assist farmers in early detection of tomato leaf diseases and in making timely management decisions, although future work is still needed to improve robustness under varying lighting conditions and to support deployment on mobile and edge devices.

References

[1] S. Jiang, C. Tao, Y. Liu, J. Zhou, dan G. Zhang, “Tomato leaf disease identification based on improved convolutional neural network,” Computers and Electronics in Agriculture, vol. 191, Art. no. 106550, 2021, doi: 10.1016/j.compag.2021.106550.

[2] S. Kamath, M. Mahale, dan S. D. Shetty, “Plant disease detection using transfer learning and CNN with real-time disease classification,” Plant Phenomics, vol. 2022, Art. no. 9832825, 2022, doi: 10.34133/2022/9832825.

[3] J. Wang, X. Zhang, Y. Zhang, dan C. Zhang, “Real-time plant disease detection system based on lightweight CNN and OpenCV,” Sensors, vol. 22, no. 15, Art. no. 5728, 2022, doi: 10.3390/s22155728.

[4] Y. LeCun, L. Bottou, Y. Bengio, dan P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.

[5] I. Goodfellow, Y. Bengio, dan A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[6] C. Shorten dan T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” Journal of Big Data, vol. 6, no. 60, 2019, doi: 10.1186/s40537-019-0197-0.

[7] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” dalam Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105, 2012.

[8] F. Chollet, Deep Learning with Python. Shelter Island, NY, USA: Manning Publications, 2018.

[9] G. Bradski dan A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol, CA, USA: O’Reilly Media, 2008.

[10] OpenCV Team, “OpenCV documentation,” 2023. [Online]. Available: https://docs.opencv.org/