Journal of Plant Sciences and Crop Protection
Volume 9 | Issue 1
2639-3336

Research Article Open Access

A Lightweight Convolutional Neural Network Framework for Automated Crop Disease Detection Using the Plant Village Dataset

Muhammad Bilal Javaid* **and Zahra Bilal**

[1*]Global Star Trading LLC, Dammam, Saudi Arabia
[2]Department of Computer Science, University of Agriculture Faisalabad, Faisalabad, Pakistan

***Corresponding Author**: Muhammad Bilal Javaid, Global Star Trading LLC, Dammam, Saudi Arabia E-mail: bilal.javaid89@gmail.com

Citation: Muhammad Bilal Javaid, Zahra Bilal (2026) A Lightweight Convolutional Neural Network Framework for Automated Crop Disease Detection Using the Plant Village Dataset J Plant Sci Crop Protec 9(1): 103

Received Date: May 12, 2026 **Accepted Date:** May 26, 2026 **Published Date:** June 04, 2026

Abstract

Crop diseases are a major cause of yield loss and reduced agricultural productivity worldwide [1]. Early and accurate identification of plant diseases is essential to prevent large scale crop damage [2]. This study presents a lightweight convolutional neural network (CNN) framework optimized for automated leaf disease detection using the Plant Village dataset, consisting of 38 classes and 54,305 RGB images. Input images were resized to 224×224 pixels and normalized to [0,1]. A stratified 70/10/20 train–validation–test split ensured balanced class representation. The model was trained using the Adam optimizer with early stopping to minimize over fitting. On the held out test set (n = 10,861), the model achieved 95.08% accuracy, with weighted precision, recall, and F1 scores of 0.951, 0.951 and 0.950, respectively, and a macro F1 of 0.931. The confusion matrix revealed strong diagonal dominance, with errors mostly among visually similar diseases such as tomato early vs. late blight. Results show that compact CNNs can deliver competitive performance while retaining low computational cost, making them suitable for mobile and real time agricultural diagnostics [6]. This work provides a reproducible baseline and supports future research involving transfer learning and evaluation under real world field conditions.

Keywords: Lightweight CNN; crop disease detection; Plant Village; image classification; deep learning; precision agriculture

Introduction

Crop diseases significantly threaten global food security by reducing productivity and increasing the cost of crop management [1]. Traditional diagnostic methods rely heavily on expert visual inspection, which is labor-intensive, slow, and prone to subjectivity [4]. Advances in deep learning, particularly convolutional neural networks (CNNs), have enabled automated plant-disease detection using leaf images [1,2].

Most existing works rely on large pre-trained architectures such as VGG, Res-Net, or Dense-Net, which, while accurate, are computationally expensive and impractical for real-time or mobile deployment [3,5]. This work addresses this challenge by presenting a lightweight CNN model that balances efficiency and accuracy, offering a reproducible baseline for future agricultural AI systems [6].

Related Work

Deep CNNs have demonstrated strong performance on plant-disease datasets. Early work by Mohanty et al. [1] showed high accuracy using AlexNet and GoogLeNet on the PlantVillage dataset. Ferentinos [2] and Too et al. [3] employed fine-tuning of deep models but highlighted issues in field generalization. Several surveys emphasize the need for reproducible pipelines and evaluation beyond controlled datasets [6,7].

Recent studies explore lightweight architectures suitable for real-time use. MobileNet-based models [8], EfficientNet variants [5], and depthwise separable CNNs show promise for deployment on mobile and embedded devices [9]. Attention-based CNNs and domain-adaptation models have also been explored [10]. However, publicly available lightweight baselines remain limited. This study fills this gap.

Materials and Methods

Dataset

The PlantVillage dataset [1] consists of 54,305 labeled RGB leaf images belonging to 38 crop-disease categories. It covers several crops such as tomato, potato, corn, apple, and grape, with both healthy and diseased samples [11].

Images were resized to 224×224 and normalized to the [0, 1] range. Stratified splitting yielded:

- Training: 38,013 images
- Validation: 5,431 images
- Test: 10,861 images

CNN Architecture

The lightweight CNN consists of the following components [12]:

- Three convolutional blocks using 3×3 kernels, ReLU activation, and max pooling
- Dropout layers (25% in convolutional, 50% in dense layer)
- A fully connected layer of 512 units

•A softmax output layer with 38 units

The model has approximately 1.9 million parameters—much smaller than deep pre-trained models such as VGG-16 or Res-Net-50 [3,5].

Training Setup

•Optimizer: Adam

•Loss: Sparse categorical cross entropy

•Batch size: 32

•Epochs: up to 20

•Early stopping with patience = 5

•Best weights restoration enabled

Results and Discussion

Training Curves

The model converged rapidly, with validation accuracy stabilizing around epoch 10 and minimal overfitting observed [13].

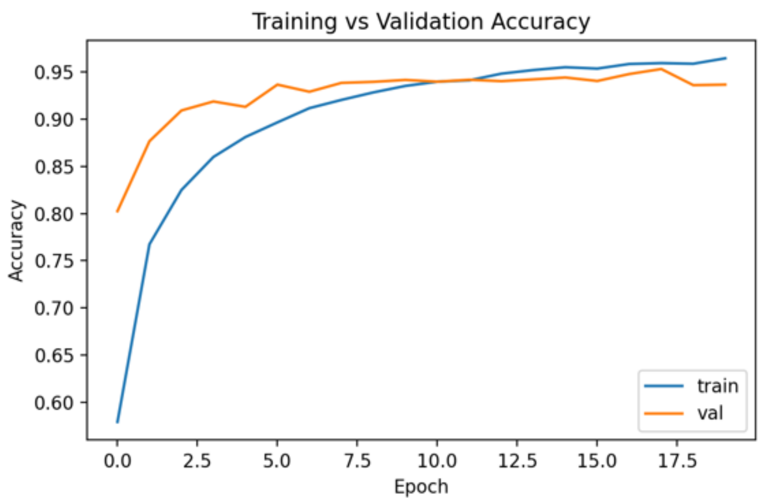

Figure 1: Training and validation accuracy across epochs.

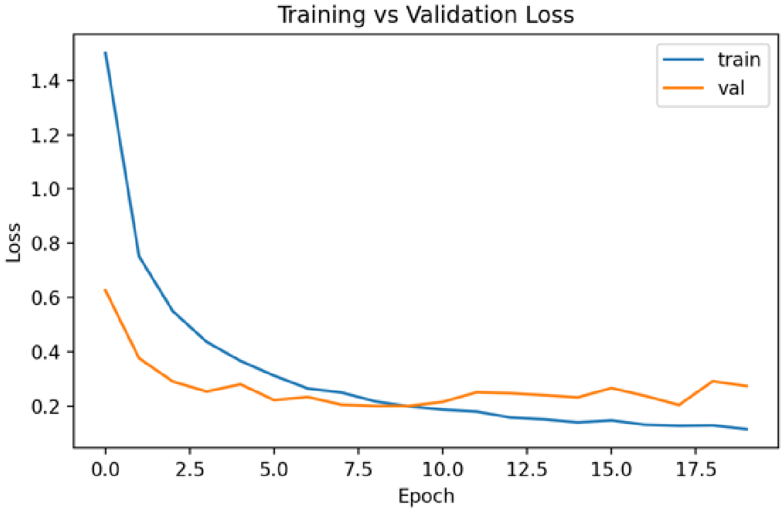

Figure 2: Training and validation loss across epochs.

Test Performance

Performance on the test set (n = 10,861):

•Accuracy: 0.9508

•Weighted precision: 0.9506

•Weighted recall: 0.9508

•Weighted F1: 0.9501

•Macro F1: 0.9313

Confusion Matrix

The confusion matrix in Figure 3 shows strong diagonal dominance, consistent with findings reported in similar studies [2,14].

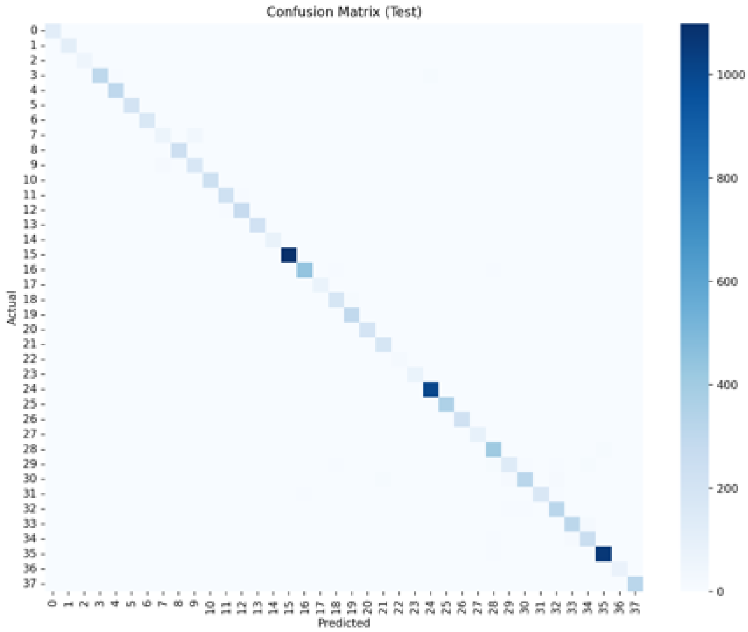

Figure 3: Confusion matrix for 38 class test set.

Error Analysis

Lowest recall classes included:

•Corn (maize) – Cercospora leaf spot / Gray leaf spot (recall = 0.61)

•Potato – healthy (recall = 0.63)

•Tomato – Early blight (recall = 0.70)

•Tomato – Late blight (recall = 0.82)

Errors often arose between visually similar diseases with overlapping color and texture patterns [15]. More varied augmentation and real-field images will improve robustness [16].Comparison with Literature

Comparison with Literature

Despite being lightweight, the model achieves performance comparable to deeper networks reported in previous works [1,2,3]. While EfficientNet and ResNet variants achieve slightly higher accuracy (>98%) [5], they require substantially more computation [17]. The proposed model offers a strong, efficient alternative suitable for edge devices and mobile deployment [6,18].

Conclusion

This study demonstrates that a lightweight CNN can effectively classify 38 plant-disease categories from the PlantVillage dataset with over 95% accuracy [1]. The model provides a computationally efficient and reproducible baseline, ideal for mobile and real-time agricultural applications [6,19]. Future work includes transfer learning [20], stronger augmentation, and evaluation on field images under variable conditions [21].

Declarations

Funding

No external funding

Conflicts of Interest

None

Ethical Approval

Not applicable

Data Availability: Plant Village dataset is publicly available.

Author Contributions

•Muhammad Bilal Javaid — Conceptualization, Methodology, Analysis, Software, Drafting, Editing

•Zahra Imtiaz — Validation, Visualization, Review & Editing

References

1. Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. Frontiers in Plant Science 7: 1419.

2. Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture 145: 311-318.

3. Too EC, Yujian L, Njuki S, Yingchun L (2019) A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture 161: 272-279.

4. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D (2016) Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience 2016: 3289801.

5. Atila Ü, Uçar M, Akyol K, Uçar E (2021) Plant leaf disease classification using Efficient-Net deep learning model. Ecological Informatics 61: 101182.

6. Kamilaris A, Prenafeta-Boldú FX (2018) Deep learning in agriculture: a survey. Computers and Electronics in Agriculture 147: 70-90.

7. Liu J, Wang X (2021) Plant diseases and pests detection based on deep learning: A review. Plant Methods 17:22.

8. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, et al. (2017) Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE International Conference on Computer Vision (ICCV) 618-626.

9. Li L, Zhang S, Wang B (2022) Plant disease detection and classification by deep learning—a review. IEEE Access 9: 56683-56698

10. Zhang X, Qiao Y, Meng F, Fan C, Zhang M (2022) Identification of maize leaf diseases using improved deep convolutional neural networks. IEEE Access 6: 30370-30377.

11. Rajput DS, Gour R, Bhatt A, Kumar Y (2023) Deep learning-based plant disease detection: a survey. International Journal of Computer Vision and Applications 13: 1-15.

12. Chen J, Liu Q, Gao L (2022) Visual tea leaf disease recognition using a CNN-based model. Sensors19: 4591.

13. Sujatha R, Chatterjee JM, Jhanjhi NZ, Brohi SN (2021) Performance of deep learning vs. machine learning in plant leaf disease detection. Microprocessors and Microsystems 80: 103615.

14. Picon A, Alvarez-Gila A, Seitz M, Ortiz-Barredo A, Echazarra J, et al. (2019) Deep convolutional neural networks for mobile plant leaf disease identification. Computers and Electronics in Agriculture161: 280-290.

15. Barbedo JGA (2019) Plant disease identification from individual lesions and spots using deep learning. Biosystems Engineering 180: 96-107.

16. Brahimi M, Boukhalfa K, Moussaoui A (2017) Deep learning for tomato diseases: Classification and symptoms visualization. Applied Artificial Intelligence 31: 299-315.

17. Haque MA, Marwaha S, Arora A, Deb CK, Mishra P, et al. (2021) Deep neural network-based screening model for COVID-19-infected patients using chest X-ray images. International Journal of Environmental Research and Public Health 18: 2845.

18. Saleem MH, Potgieter J, Arif KM (2019) Plant disease detection and classification by deep learning. Plants 8: 468.

19. Ma J, Du K, Zheng F, Zhang L, Gong Z, et al. (2020) A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. IEEE Access 7: 179135-179146.

20. Singh AK, Ganapathysubramanian B, Sarkar S, Singh A (2018) Deep learning for plant stress phenotyping: Trends and future perspectives. Trends in Plant Science 23: 883-898.

21. Kaur P, Harnal S, Tiwari R (2022) Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction Sensors 22: 575.

22. Rauf HT, Saleem BA, Lali MIU, Khan MA, Sharif M, et al. (2019) A citrus disease detection and classification using optimized weighted segmentation and feature selection based on machine learning. Applied Soft Computing 83:105638.

23. Agarwal M, Singh A, Arjaria S, Sinha A, Gupta S (2022) ToLeD: Tomato leaf disease detection using convolution neural network. Procedia Computer Science 167: 293-301.

24. Pandian JA, Kumar VD, Geman O, Hnatiuc M, Arif M, et al. (2021) Plant disease detection using deep convolutional neural network. Applied Sciences 12: 6982.

25. Chouhan SS, Singh UP, Jain S (2020) Applications of computer vision in plant pathology: A survey. Archives of Computational Methods in Engineering 27: 611-632.