Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Deep Learning Based Tomato Leaves Disease Detection: A Comprehensive Approach Using Convolutional Neural Networks

Krishna Kumar *[1], Mohan Vishal Gupta [2], Amit Kumar[3]

Ashish Bishnoi[4], Manish Joshi[5]

[1]Professor, Institute of Technology & Management, GIDA, Gorakhpur, U.P., India.

[2,3,4,5] Associate Professor, CCSIT, Teerthanker Mahaveer University, Moradabad, India.

*__Corresponding Author Email__: kk_gkp@rediff.com

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Diseases of tomato leaves have a high impact on the production of crop and the yield can be reduced if the diseases are not identified at primary stage. The traditional way of detecting diseases by human experts is slow and laborious, and also less accurate in many cases, particularly for large farms. To address this issue, the paper proposes a deep learning based tomato leaf disease detection system utilizing the VGG19 model. The Plant Village dataset, which has images of healthy and infected tomato leaves, was utilized for both training and testing. Prior to training, images were resized, normalized and augmented with rotation, flipping and scaling to enhance model robustness and performance in various testing conditions. This model is trained on 80-10-10 split for training, validating and testing. To enhance learning speed, the categorical cross-entropy loss function and the Adam optimizer were employed. Results of experiments indicate the proposed VGG19 based model is superior enough, achieving the validation accuracy of 91.56% and the test accuracy of 98.36%. Training and validation plot illustrate the model learns well with little overfitting. The low test loss of 0.07 also suggests a good generalization on unseen data. This work demonstrates that deep learning can be used to generate fast, accurate and reliable solutions for tomato leaf disease detection. The proposed technique can aid the farmers in several ways for example detection of disease, preventing the damage to the crop, thereby enhancing the agricultural production. In addition, the work provides a robust platform for developing real time applications on mobile and IoT platforms.

Keywords: Tomato, Leaves Disease, VGG, CNN architecture, Early detection

1. Introduction

Agriculture is a key industry in both economic and human terms, as it supplies the food and raw materials necessary for survival and a stable economy. All that's changed is that you're presenting more evidence to support the claim: Plant disease epidemics are also being intensified by climate

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

change, and these threaten global food supplies and economies [1]. It is estimated that crop yields have been decreased by 20 to 40% by plant diseases, with billions of dollars of economic losses every year, according to the World Bank. In countries such as India where agriculture is a critical factor for national development, the menace of crop diseases is a grave problem undermining the livelihood of millions of farmers and feeding a population that knows no bounds [2]. The traditional methods for monitoring and managing disease in crops are visual and subjective assessments in the field by experienced agronomists. Although these techniques are useful, they require a lot of work, take a lot of time and are sometimes subject to human error. Apart from that, in a world with a growing population demanding more and more food and with agriculture under increasing strain due to climate change, there is a pressing need for more efficient, more accurate, and more scalable methods for detection of disease.

Advances in artificial intelligence (AI) and computer vision, in particular Convolutional Neural Networks (CNNs), have transformed the field of image processing and object recognition in recent years [3][4]. Being a class of deep learning models, CNNs have shown excellent ability in learning complex visual spatial information, which makes them also applicable in plant disease detection. In this paper, a full investigation on the use of CNNs for plant disease detection is presented. It addresses precision disease recognition in agriculture, demonstrates challenges in farmers and agriculture systems, and illustrates how CNN-based models can revolutionize disease diagnosis, allow early treatment, and safeguard crop health and food security [5][6]. The paper presents a review of different CNN architectures/models, datasets, strategies/techniques for testing and applications to multiple plant diseases. It reviews the performance, challenges and advancements of CNN based disease detection methodologies and solutions for the transformation the farming systems to feed next generation. More generally, this may be considered as a selective critical review on how transformative the role of CNN has been to plant disease detection, and what emerging research trends suggest for increasingly adept, accurate, and scalable means for reducing the burden imposed by crop diseases on agricultural productivity and food security [11] [12].

2. Literature Review

Livelihoods are maintained, and billions of people are fed. In countries such as India, it is of crucial importance for prenatal growth and development. However, it faces many challenges, including plant diseases, which endanger food production, food security and the economy. Computer vision has witnessed dramatic advances in recent years especially with deep learning based approaches such as Convolutional Neural Networks (CNNs) which have become essential tools for image processing and have found a wide range of applications in different fields, e.g. medicine, smart cars, and agriculture. Some researches were conducted on the topic of employing CNNs for detecting plant leaf diseases, taking advantage of the fact that they can learn complex features from images [7][8].

Learning approach based on ResNet-34 showing that it can competently classify plant leaf images with special features. The ResNet- 34 model was trained with a large number of labeled plant leaf images and it attained high accuracy in classifying different types of leaf diseases and abnormalities. This culminated in the development of a web tool that enables users to identify plant diseases or pest attacks by submitting images of leaves [9] [10].

In a different approach, employed Multilayered Convolutional Neural Networks, achieving an impressive accuracy of 98.5% in classifying leaf diseases by utilizing the "New Plant Diseases Dataset" from Kaggle and incorporating the Histogram of Oriented Gradients (HOG) feature descriptor for feature extraction. Similarly, authors utilized ConvNets for plant disease detection, attaining accuracies of 98.3%, 98.5%, and 95% for detecting diseases in potato, pepper, and tomato plants, respectively, using the Plant Village dataset from Kaggle. Eenhanced plant leaf disease identification by employing Convolutional Neural Networks trained on more than 39 classes of plant leaf diseases and background images. Employing six data augmentation methods, including gamma correction, image flipping, PCA color augmentation, rotation, noise injection, and scaling, augmented model performance, achieving a 96.46% classification accuracy, surpassing transfer learning approaches [13] [14]. colleagues deployed a convolutional neural network for plant disease detection, successfully classifying 12 diseases with an 88.80% accuracy rate using a dataset of 3,000 high-resolution RGB images. However, challenges such as computational intensity and a relatively higher number of false negative predictions, indicated by a low F1 score of 0.12, were noted.

Overall, these studies demonstrate the potential of CNNs in revolutionizing plant disease detection, paving the way for more efficient and accurate agricultural practices.

2.1 Research Gap

There are still some significant problems despite the deep learning and CNN models being utilized for tomato leaf disease detection in the majority of works. Most research relies on clean, lab-based datasets such as Plant Village, which do not represent real farm environments that include variable lighting, background noise, and leaf overlap. This sets a difference between the experimental results and real situation application. Moreover, shallow architectures or simple CNN models are often employed which restricts the capacity to capture subtle discriminative patterns of disease among visually similar infections of tomato leaves. Also, in some studies the preprocessing and augmentation approaches are weak which decreases the robustness of the model. Another significant gap is the scant attention to deployment, with only a handful of models being optimized for mobile, IoT devices, or real-time detection at the field level. In addition, validation loss, convergence performance, and generalization performance are rarely evaluated in a comprehensive manner.

3. Proposed Methodology

A neural network is modeled on the visual cortex of a cat. A special feature of CNNs is pattern recognition which is very good in finding patterns in all types of images. CNNs are hierarchical and have multiple layers. Input, output, pooling, normalization, convolution and fully connected are some of the layers used in CNN. By going up and down these levels it is possible to understand the complex principles underlying the revolutionary power of the CNN to decipher complex visual information.

A. **Input layer:** For an image processing application the network is fed with raw input image data at this layer. Every pixel of the image is considered as a neuron to this layer. Secondly, convolutional layers. Convolutional layers are the building blocks of convolutional neural networks (CNNs), and they play a key role in obtaining meaningful features from input data, especially for tasks related to the processing of images. The convolution operation is the heart of convolutional layers. This method the input data, usually an image is overlapped with a small filter known as a kernel. What a filter does is that it detects certain patterns or features (such as edges, textures or even more complex structures). The filter generates a new matrix called the feature map by sliding over the input and applying the element-wise multiplication followed by the sum. Each element of the feature map represents activation of a neuron which means the presence of a particular feature in the input. Then, classify the image using the features of the convolutional layer after the features have been obtained.

B. **Activation Function:** To add nonlinearity, an activation function (often ReLU, or Rectified Linear Unit) is applied element-by-element following the convolution procedure. The network can learn intricate linkages and representations thanks to its nonlinearity. Here is how the function (ReLU) is defined.

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

C. **Pooling layer –** By lowering the feature map's spatial dimension, the pooling layer down sampling. For instance, max-pooling reduces computational complexity while maintaining the most prominent features by choosing the maximum value from a collection of nearby pixels.

D. **Fully Connected Layer** – High-level properties from the preceding layer are connected to neurons in a fully connected layer, and their weights are modified during training. The learnt features are integrated at this layer for the final output, or classification. Every neuron in a completely linked layer is coupled to every other neuron in the layer above it.

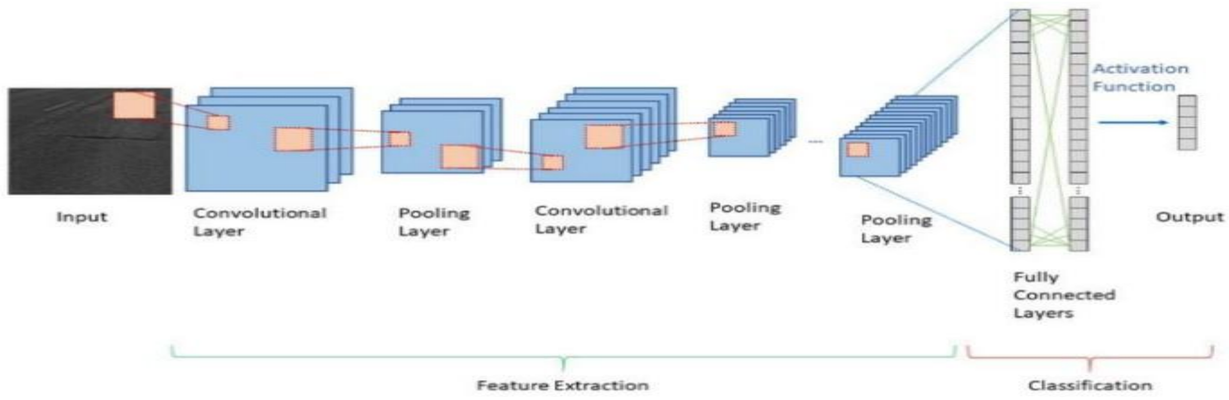

Figure 1. Convolutional Neural Network Architecture

This Figure 1, CNN can learn to recognize significant visual features in tomato leaf images, such as spots, texture changes and infected regions, without any manual pre-processing, through the convolution and pooling layers. Using a soft max output, the leaf is classified as either healthy or a specific disease category by the fully connected layers.

3.1 Dataset description

The study's dataset comes from KAGGLE's collection, "Plant Village" This dataset, which comprises an extensive collection of images of various types of diseases which affect tomato leaves, was created especially for studies on tomato plant disease identification. The two primary files, "Valid" and "Train," each comprising a distinct category of disease images like Bacterial Spot, Early Blight, Late Blight, comprise the datasets used in this study. This arrangement makes it possible to identify clearly between various diseases that are found in tomato leaves, enabling focused examination of characteristics linked to the disease's presence. This is how the dataset appears. Figure 2. Plant Disease images dataset

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

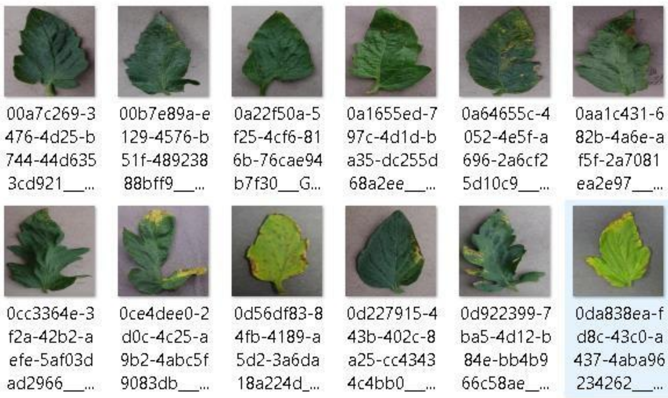

Figure 2. Plant Disease images dataset

3.2 Preprocessing

With the aim of ensuring consistent shape and simplifying subsequent analysis, the datasets were modestly scaled after collection. then a standardization procedure was applied to normalize the pixel intensities to improve the comparability of the images and to reduce the potential difference caused by different acquisition settings. To further augment the dataset size and enhance model robustness, more transformations were applied such as random rotation, reflection, and scaling operators. The expansion approach aims to enhance the generalisation ability of the model to different input conditions by presenting the model with more diverse scenario. When combined, the processes of augmentation, normalization, and resizing result in a sufficiently well-prepared data that can better accommodate variations in image characteristics and contents and that can serve as a solid foundation for training and testing plant disease detection models contributed to the repository.

3.3 Model Training and Validation

In order to develop the Plant disease detection system, the model was trained using by CNN. The training set, validation set and test set are the three subsets that were obtained from the dataset. 80% of the data are dedicated to training the CNN model, 10% to validating the model during training, and the last 10% to performing an independent evaluation during testing. This is referred to as an 80-10-10 division. It was final. During training, the CNN parameters were refined iteratively by a certain loss function—categorical crossentropy in most cases—for multi class classification problems such as disease detection. For more efficient gradient-based updates, the Adam optimizer was used, and overfitting signs were captured by monitoring model performance on the validation set at regular intervals.

A series of convolutional layers for hierarchical feature extraction, followed by a pooling layer to achieve spatial down sampling, are employed in the architecture of the CNN. To improve the generalization of the model and prevent overfitting, dropout and batch normalization layers were strategically added along with full connected layers for the classification. The model has been changed to be nonlinear by the application of the ReLU (Rectified Linear Unit) activation function. To evaluate the performance of the trained CNN to detect tomato leaves disease, validation of the proposed CNN was done on a separate test set using such parameters as Training accuracy and Validation accuracy, training loss and validation loss.

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

4. Results and Discussion

Promising results from the performance evaluation of our Tomato leaves disease detection system's best model (using VGG19) showed the system's effectiveness and robustness. Acknowledged for having outstanding validation accuracy, the model proved to be exceptionally accurate with 98.36% on an unprecedented test set. This accuracy demonstrates the model's dependability in practical applications and shows how effectively it generalizes to fresh, balanced data. With an accuracy of 98% on the test set and 91% on the validation set, the comprehensive performance table demonstrates consistently good accuracy on both sets of data. All things considered, our findings validate the potential of our CNN-based tomato leaves disease identification system and establish it as a viable instrument for early identification of leaves disease. Precise and effective medical image processing in clinical environments.

We give the visualization of key performance indicators, namely accuracy and loss, throughout the training process in order to thoroughly assess the effectiveness of our suggested Tomato leave disease detection system. By revealing how well convolutional neural network (CNN) models are able to learn from and adjust to the complexity of the training input, these graphs play a crucial role as diagnostic tools.

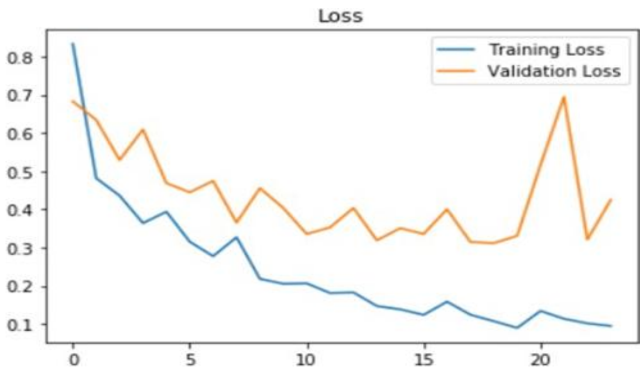

Figure 3. Model Loss visualization

Figure 4. Model Accuracy visualization

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Table 1: Performance table of the model

	Validation set	Test set
Accuracy	91.56%	98.36%
Loss	0.73	0.07

The proposed model was trained and tested on the datasets, and the performance indicates that the model has strong learning ability and can achieve high accuracy. Model Loss curve is shown in Figure 4, loss progressively decreases as epoch increases, which means that the prediction errors are gradually reduced with the optimization of model. The validation loss is down to 0.73 which means the model manages to generalize its learning without massive overfitting. Figure 4 shows the model accuracy curve that follows a similar trend as that of model loss with a monotonic increasing epochs. It turns out that the validation accuracy converges to 91.56%, which indicates a good classification performance in training. The smooth shape of these curves shows that the model learns consistently well throughout training.

To provide additional evidence for generalization, we also tested the model on data that had not been seen before. The test results exceed validation ones in the accuracy of 98.36% with an even lower loss of 0.07 which is amazing. This shows that the model can be used to predict with high accuracy and stability even in practical applications.

In summary, the training loss, accuracy curves and the performance measures indicate that the model training is reliable, robust and the model can achieve good predictive performance on the test set.

5. Conclusion

In this paper, we have proposed a tomato leaf disease detection model based on VGG19. The objective was to develop a robust and novel approach for detection of various tomato leaf diseases using images. The Plant Village dataset was rigorously preprocessed (resizing, normalization and heavy augmentation) so that it can achieve high level of generalization. The model was trained on 80-10-10 split using Adam optimizer along with the loss function categorical cross-entropy. It is obvious from this testing results that the proposed method is effective. The model also achieved a very good validation accuracy of 91.56% and an outstanding test accuracy of 98.36%, which indicates the model can classify tomato leaf diseases with high dependability. The loss curves had a smooth convergence while the accuracy curves revealed a stable learning during the entire training period. A test set performance at a very low loss of 0.07 demonstrates that the model generalizes well to unseen data. In summary, the results clearly indicate the promise of deep learning (especially VGG based models) for accurate plant disease detection. The approach developed in this study lays a solid base for a real time disease monitoring system which can support farmers in decision making in timely fashion. With a minimum amount of human input, early diagnosis, and better disease control, the proposed framework may positively impact crop health, yield, and later food security globally. This work is a significant contribution in the direction of applying AI in precision agriculture for intelligent and sustainable farming.

Future Work:

Potential future work includes evaluating the model on real-field adding multispectral or contextual data, and developing mobile or IoT based application for online detection. Generalizing the system to other crops as well as applying sophisticated transfer learning can increase the accuracy, speed and practical value of the system.

Author Contributions

All of the writers agreed on the study's substance. The author has reviewed and approved the final manuscript.

Funding

For this work, the authors did not receive any special funding.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this work.

Data Set:
https://www.kaggle.com/datasets/charuchaudhry/plantvillage-tomato-leaf-dataset

References

[1] Sowmya, B., Guruprasad, S. Deep learning based plant health disease detection in tomatoes using inception v4 convolutional neural network and YOLO V8. DiscovArtifIntell 5, 278 (2025). https://doi.org/10.1007/s44163-025-00540-1.

[2] T. Vijayakumar, S. Santhi and M. Gokila, "A Deep Learning - based Approach for Tomato Plant Leaf Disease Detection and Classification," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1251-1255, doi: 10.1109/ICOSEC54921.2022.9952007.

[3] Abdullah, A.; Amran, G.A.; Tahmid, S.M.A.; Alabrah, A.; AL-Bakhrani, A.A.; Ali, A. A Deep-Learning-Based Model for the Detection of Diseased Tomato Leaves. Agronomy 2024, 14, 1593. https://doi.org/10.3390/agronomy14071593.

[4] Sharma, J., Al-Huqail, A.A., Almogren, A. et al. Deep learning based ensemble model for accurate tomato leaf disease classification by leveraging ResNet50 and MobileNetV2 architectures. Sci Rep 15, 13904 (2025). https://doi.org/10.1038/s41598-025-98015-x

[5] TabiaTanzinPrama and Mangsura Kabir Oni. 2025. Optimized Custom CNN for Real-Time Tomato Leaf Disease Detection. In Computational Science – ICCS 2025 Workshops: 25th International Conference, Singapore, Singapore, July 7–9, 2025, Proceedings, Part III. Springer-Verlag, Berlin, Heidelberg, 81–88. https://doi.org/10.1007/978-3-031-97564-6_6

[6] Ghosh, H., Rahat, I.S., Emon, M.M.R. et al. Advanced neural network architectures for tomato leaf disease diagnosis in precision agriculture. Discov Sustain 6, 312 (2025). https://doi.org/10.1007/s43621-025-01149-1

[7] Ullah, Z.; Alsubaie, N.; Jamjoom, M.; Alajmani, S.H.; Saleem, F. EffiMob-Net: A Deep Learning-Based Hybrid Model for Detection and Identification of Tomato Diseases Using Leaf Images. *Agriculture* **2023**, *13*, 737. https://doi.org/10.3390/agriculture13030737

[8] Guerrero-Ibañez, A.; Reyes-Muñoz, A. Monitoring Tomato Leaf Disease through Convolutional Neural Networks. *Electronics* **2023**, *12*, 229. https://doi.org/10.3390/electronics12010229

Received: 2026-02-04
Accepted: 2026-03-30
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30203

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[9] A. Bhilare, D. Swain, and N. Patel, "Comparative Analysis of Deep Learning Models for Accurate Detection of Plant Diseases: A Comprehensive Survey", EAI Endorsed Trans IoT, vol. 10, Dec. 2023, doi: 10.4108/eetiot.4595.

[10] S. Senthil Pandi, P. Kumar, R. M. Suchindhar and R. Karthick, "Hybrid Multistage System for Plant Disease Classification and Treatment Recommendation," 2025 8th International Conference on Circuit, Power & Computing Technologies (ICCPCT), Kollam, India, 2025, pp. 607-612, doi: 10.1109/ICCPCT65132.2025.11176740.

[11] D. R. Thummar, "Enhanced Tomato Leaf Disease Detection using YOLOv11: A Deep Learning-based Approach," 2025 5th International Conference on Soft Computing for Security Applications (ICSCSA), Salem, India, 2025, pp. 847-851, doi: 10.1109/ICSCSA66339.2025.11171413.

[12] S. P. Patnayakuni, "Tomato: Different Leaf Disease Detection Using Transfer Learning Based Network," in Journal of Mobile Multimedia, vol. 18, no. 3, pp. 743-756, May 2022, doi: 10.13052/jmm1550-4646.18313

[13] "13 Hybrid Leaf Generative Adversarial Networks Scheme for Classification of Tomato Leaves—Early Blight Disease Healthy," in Advancement of Deep Learning and its Applications in Object Detection and Recognition, River Publishers, 2022, pp.261-282.

[14] J. H. Kamdar; M. D. Jasani; J. D. Jasani; Jeba Praba; John J. Georrge, "11 Artificial intelligence for plant disease detection: past, present, and future," in Internet of Things and Machine Learning in Agriculture: Technological Impacts and Challenges, De Gruyter, 2021, pp.223-238.

[15] L. K. Ndovie and E. Masabo, "Leveraging MobileNetV3 for in-field tomato disease detection in Malawi via CNN," SAIEE Africa Research Journal, vol. 115, no. 3, pp. 74–85, Sep. 2024, doi: 10.23919/SAIEE.2024.10551304.

[16] H. E. David, K. Ramalakshmi, H. Gunasekaran, and R. Venkatesan, "Literature review of disease detection in tomato leaf using deep learning techniques," in Proc. 7th Int. Conf. Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 274–278, doi: 10.1109/ICACCS51430.2021.9441714.

[17] B. Khan, S. Das, N. S. Fahim, et al., "Bayesian optimized multimodal deep hybrid learning approach for tomato leaf disease classification," Scientific Reports, vol. 14, Art. no. 21525, 2024, doi: 10.1038/s41598-024-72237-x.

[18] O. Attallah, "Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection," Horticulturae, vol. 9, no. 2, Art. no. 149, 2023, doi: 10.3390/horticulturae9020149.

[19] X. Guo, "Automatic detection of tomato leaf disease using an adopted deep learning algorithm," Journal of Intelligent & Fuzzy Systems, vol. 46, no. 4, pp. 7909–7921, 2024, doi: 10.3233/JIFS-236905.

[20] D. Karimanzira, "Context-aware tomato leaf disease detection using deep learning in an operational framework," Electronics, vol. 14, no. 4, Art. no. 661, 2025, doi: 10.3390/electronics14040661.