Published by :
http://www.ijert.org
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 10, October - 2025

AI-Based Tomato Plant Disease Detection using CNN

Ms. Shinde Megha Bajirao
PG – Computer Engineering Student
Department of Computer Engineering, SND College of Engineering and Research Center Yeola,
Savitribai Phule Pune University
Dist Nashik , MS India

Prof. Kanade Poonam Ganpat
Asst. Professor
Department of Computer Engineering, SND College of Engineering and Research Center Yeola ,
Savitribai Phule Pune University
Dist Nashik , MS India

Abstract:
In the context of fostering a robust agricultural economy, particularly in nations like India, the early detection of plant leaf diseases stands as a pivotal requirement. Beyond its implications for agricultural productivity, the significance of timely disease detection extends to ensuring the nutritional integrity of a sizable population. This survey paper proposes a comprehensive exploration of methodologies for the early identification of tomato plant leaf diseases, emphasizing the integration of advanced image processing techniques, convolutional neural networks (CNNs), and open-source algorithms. By synthesizing and evaluating existing research in this domain, the paper aims to provide an in-depth understanding of the state-of-the-art technologies, their strengths, limitations, and potential areas for improvement. The culmination of this survey contributes to the development of a dependable, secure, and precise framework tailored to the specificities of tomato plant diseases. The insights derived from this survey are poised to inform and guide future research endeavors, offering a holistic perspective on the advancements in early disease detection and predictive mechanisms within the realm of agricultural practices.

Keywords:

Convolutional Neural Networks (CNNs), Open-Source Algorithms, Early Disease Detection, Plant Leaf Diseases.

1. INTRODUCTION

Tomatoes, recognized globally for their versatility, are a staple crop in India, contributing significantly to agricultural output. With an annual production exceeding 232 million tonnes, a substantial portion is allocated for processing, making tomatoes a vital commodity in the Indian food industry. Beyond economic significance, tomatoes offer essential health benefits, containing antioxidants like lycopene, folate, potassium, vitamin K, and vitamin C.

In the Indian agricultural landscape, the cultivation of tomatoes is widespread, providing sustenance and livelihoods, particularly during the winter season. Despite its importance, tomato farmers face persistent challenges, notably the prevalence of diseases impacting tomato plant leaves. Widespread crop concerns like rust infestation, aphid damage, and blight represent considerable challenges to agricultural productivity. Timely intervention is crucial, especially when diseases manifest during critical stages of production.

This research addresses the imperative of early disease detection by exploring a technological solution anchored in machine learning. Manual prediction methods are often impractical for farmers due to time constraints and technical complexities. In contrast, a computerized system utilizing machine learning techniques holds promise for efficiently monitoring plant health, growth status, and predicting diseases early in the production cycle.

A focal point of this study is the implementation of Convolutional Neural Networks (CNNs), a potent deep learning algorithm designed for image identification and analysis. CNNs offer advantages, requiring minimal preprocessing compared to conventional algorithms, enabling the neural network to autonomously glean insights from input images. The versatility of CNNs in image-related operations, including classification and recognition,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 10, October - 2025

positions them as a compelling choice for predicting tomato leaf diseases in the Indian agricultural context.
This research aims to develop an intelligent machine learning model tailored for the efficient, cost-effective, and accurate prediction of tomato leaf diseases in India. By leveraging the capabilities of CNNs and machine learning, this study aspires to contribute to the resilience of tomato farming in India, offering a technology-driven solution to enhance crop health and overall production in the Indian agricultural landscape.

2. BACKGROUND

Tomato plants (Solanum lycopersicum) have long been recognized as one of the most economically and nutritionally important crops worldwide. Originating in the Andes region of South America, tomatoes have undergone centuries of cultivation and adaptation to various climates, becoming a staple in diverse cuisines. In addition to their culinary significance, tomatoes are valued for their nutritional content, comprising essential vitamins and antioxidants. The adaptability of tomato plants to different growing conditions has led to their widespread cultivation, making them a crucial component of global agriculture.
Tomatoes are sensitive to frost and require warm temperatures to thrive. Tomatoes need consistent moisture throughout the growing season, especially during flowering and fruiting stages. However, they are susceptible to waterlogged conditions, so well-drained soil is crucial.
Tomato crops are susceptible to a variety of diseases and pests that can significantly impact yield and overall plant health. Timely and precise identification of these concerns is critical for deploying prompt interventions and mitigating the economic and nutritional setbacks linked to impaired crops.

I. Categories of Diseases Found on Tomato Leaves:

Common tomato leaf diseases can be categorized based on the pathogens or causes that lead to their occurrence. Here are some categories of common tomato leaf diseases:

1. Fungal Diseases:

- Early Blight: Caused by the fungus Alternaria solani, early blight manifests as dark brown lesions with concentric rings on older leaves. It can lead to premature defoliation, reducing the plant's photosynthetic capacity and overall vigor.
- Late Blight: Caused by the oomycete Phytophthora infestans, late blight results in dark lesions with a water-soaked appearance, leading to rapid defoliation. The disease can devastate entire tomato crops, affecting plant health and yield.
- Septoria Leaf Spot: Caused by the fungus Septoria lycopersici, this disease presents as small, circular spots with dark borders on lower leaves. While it doesn't usually lead to severe defoliation, it can affect overall plant health and reduce yield.

2. Bacterial Diseases:

- Bacterial Spot: Caused by Xanthomonas campestris pv. vesicatoria, bacterial spot causes dark lesions with a water-soaked appearance on leaves, fruits, and stems. Severe infections can lead to defoliation, affecting plant health.
- Bacterial Speck: Caused by Pseudomonas syringae pv. tomato, bacterial speck is characterized by small, dark lesions with a speckled appearance on leaves and fruit. While it may not lead to extensive defoliation, it can impact fruit quality and yield.

3. Viral Diseases:

- Tomato Mosaic Virus (ToMV): ToMV causes mosaic patterns, curling, and distortion of leaves. Infected plants may exhibit stunted growth and reduced yield due to disrupted physiological processes.
- Tomato Yellow Leaf Curl Virus (TYLCV): TYLCV causes yellowing, curling, and stunted growth of leaves. Infected plants often exhibit reduced vigor and diminished fruit production.

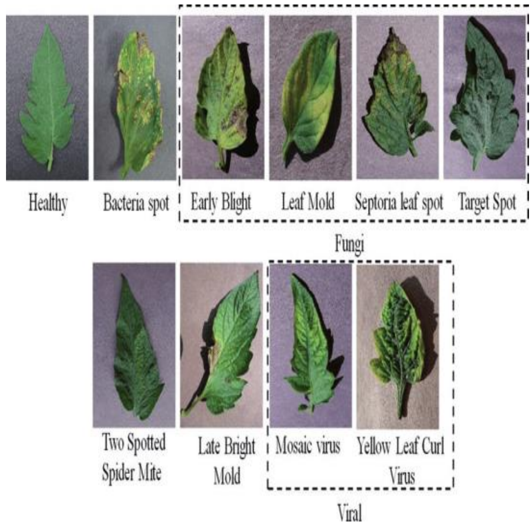

Fig 1: Images of tomato leaves with various diseases

II. TRADITIONAL METHODS FOR DISEASE DETECTION

Traditional methods for tomato plant leaf disease detection have historically relied on manual techniques such as visual inspection, manual observation, and symptom recognition. Visual inspection involves physically examining the plants for any visible signs of disease, including discoloration, lesions, or abnormal growth patterns. Manual observation entails closely monitoring the plants over time, looking for changes in leaf color, shape, or overall plant health. Symptom recognition involves identifying characteristic signs of specific diseases based on visual cues.
However, these traditional methods come with inherent limitations and challenges. Visual inspection and manual observation are highly subjective and dependent on the expertise of the observer. The accuracy of disease identification may vary among individuals, leading to inconsistencies in diagnosis. Additionally, these methods are time-consuming, especially in large agricultural settings, making it impractical for early and efficient detection.

Symptom recognition is hindered by the fact that many plant diseases exhibit similar visual cues, leading to misdiagnosis or delayed response. Furthermore, these traditional approaches are often limited to detecting visible symptoms, and by the time symptoms manifest, the disease may have already progressed to an advanced stage, making it challenging to implement effective control measures.

Moreover, human error, fatigue, and environmental conditions can impact the reliability of traditional methods. The lack of quantitative data and standardized procedures further contribute to the challenges associated with these techniques. As a result, there is a pressing need for more advanced and technology-driven approaches,to overcome the limitations inherent in traditional methodologies and boost the accuracy and efficiency of detecting diseases in tomato plant leaves, a shift towards modern technologies is imperative. Leveraging advancements such as image processing and machine learning proves to be a promising avenue to tackle these challenges. These technologies offer robust and timely solutions for the identification of diseases in tomato crops.

3. ELATED WORK

In this scenario, the authors introduced an model for the classification of nine distinct diseases impacting tomato crops in the PlantVillage dataset. Employing a lightweight MobileNetV2 architecture, the proposed approach achieved an impressive peak accuracy of 99.30% when tested on the PlantVillage dataset.(1).

Researchers in this study employed a streamlined CNN design for the classification of two specific diseases affecting tomato crops as documented in the PlantVillage dataset, which encompasses 5225 images. The proposed methodology demonstrated exceptional performance, attaining a peak accuracy of 99.2% when applied to the comprehensive PlantVillage dataset. (2)

In this research, the investigators utilized a concise LSFCNN & ACMRCNN architecture for the identification of two distinct diseases affecting tomato crops within the AI CHALLENGER dataset, which consists of 14,817 images. The proposed method demonstrated exceptional performance, reaching a peak accuracy of 99.30% when applied to both the AI CHALLENGER dataset and the mentioned collection of images. Additionally, the study conducted a thorough literature review to situate and enhance the understanding of the existing knowledge in the field.(3)

In recent times, Convolutional Neural Networks (CNNs) have become widely popular for discerning plant diseases, particularly in the realm of pinpointing issues in tomato plants. This technology empowers the accurate classification of extensive datasets with notable precision. Researchers have introduced multiple CNN variations specifically customized for diagnosing diseases in tomato plants, achieving an impressive detection accuracy of 99.53% in certain scenarios. 4; 5; 6

In addressing the challenge of disease recognition in tomato crops, a concise CNN model incorporating eight hidden layers was introduced. The researchers harnessed a lightweight CNN architecture to distinguish among nine specific diseases commonly observed in tomato crops through the analysis of the PlantVillage dataset. Impressively, the proposed method demonstrated its effectiveness by reaching a peak accuracy of 98.4% on the PlantVillage dataset. (7)

A recent investigation revealed that utilizing a deep learning ResNet50 architecture on the tomato dataset enables the early-stage detection of previously unseen diseases in tomato plants with a maximum accuracy of 97%, particularly when applied to color images..(8)

Industry demands rapid application prototyping and simulations of the protocols to save time, money and energy. Use of intelligent systems with high-speed communication, the sector is moving towards the standardizations. (9)

4. PROPOSED METHODOLOGY

The paper aims to present a methodology for effectively classifying tomato leaf diseases and recommending optimal solutions. This approach leverages image processing techniques and state-of-the-art algorithms, implemented through the open-source programming language Python. Illustrated in Figure 2 is a flow diagram outlining the proposed method. To ensure accuracy, it is crucial that the image is preferably captured directly from the plant. User-uploaded images for detection cannot be directly compared to dataset images, as this could potentially mislead the recognition system. The process involves subjecting images to a sequence of feature extraction mechanisms and subsequent segmentation to identify the affected portion of the leaf, facilitated by the CNN Classifier. The obtained results are then employed to detect and classify tomato plant diseases, offering farmers suitable precautionary measures. The key steps encompass pre-processing and feature extraction for both the input image and dataset images, contributing to an enhanced level of prediction accuracy.

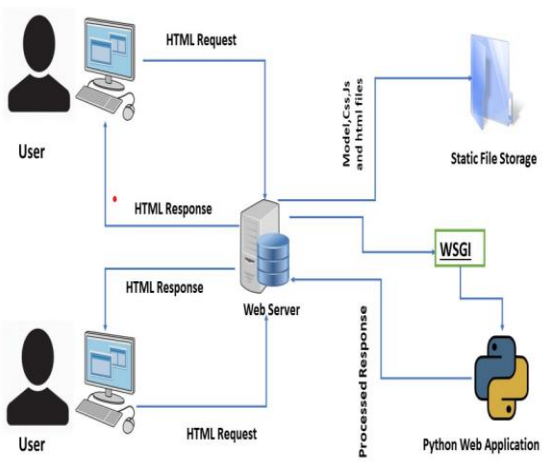

Fig 2: System Architecture

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 10, October - 2025

5. ODULE ARCHITECTURE

The Convolutional Neural Networks (CNNs) presented in this system feature a two-layer architecture comprising convolutional and pooling layers. The model introduces a subsampling layer that encompasses three sets of convolutional layers, each followed by corresponding pooling layers. This pattern is iterated three times before the resulting output feature maps undergo flattening operations. The weights assigned during this process contribute to the fully connected layer, aiding in interpretation and ultimately facilitating the prediction of diseases.

I. Dataset Acquisition

Comprising both healthy and affected instances, the dataset comprises 18,160 images capturing various conditions of tomato leaves., was sourced from the Kaggle data science website. To facilitate model training and evaluation, the dataset was segregated into training and validation subsets, with 90% (8,175 images) assigned for training and the remaining 10% (908 images) reserved for validation. This division ensured a representative assessment of the model's generalization capabilities. (10)

Table 1. Dataset

Total Data	Training and Validation Data	Test Data
18160	9083	9077

II. ATASET DESCRIPTION

Employing the imageDataGenerator() function from the TensorFlow library, the data was further divided into three essential blocks: Train, Test, and Validation. Post data splitting, each disease class was assigned 9077 images for testing and 9083 images for training and for validation. This meticulous distribution is crucial for ensuring robust model training, thorough testing, and reliable validation.

This architectural framework, coupled with a comprehensive dataset, forms the foundation for the model's training and evaluation, contributing to its effectiveness in predicting tomato leaf diseases.

6. ONCLUSION

This paper delves into the essential need for timely detection of crop diseases, specifically within India's prominent agricultural sector, with tomatoes as the focal point. Given the significant reliance of the population on agriculture, the objective is to identify and categorize ten distinct diseases affecting tomato crops. The proposed methodology introduces a streamlined convolutional neural network model, tailored for efficiently classifying illnesses in tomato leaves. Despite its simplicity, the model showcases a high level of accuracy in disease categorization, positioning itself as a valuable decision-making tool for farmers. The primary emphasis is on quick and reliable disease diagnosis, achieved with minimal computational effort, with the overarching goal of empowering farmers and contributing to the overall health and productivity of the agricultural sector.

REFERENCES

[1] S. Ahmed, M. B. Hasan, T. Ahmed, M. R. Sony, and M. H. Kabir, “Less is more: Lighter and faster deep neural architecture for tomato leaf disease classification,” IEEE Access, vol. 10, p. 68868–68884, 2022.

[2] A. Saeed, A. A. Abdel-Aziz, A. Mossad, M. A. Abdelhamid, A. Y. Alkhaled, and M. Mayhoub, “Smart detection of tomato leaf diseases using transfer learning-based convolutional neural networks,” Agriculture, vol. 13, no. 1, p. 139, 2023.

[3] R. Raja Kumar, J. Athimoolam, A. Appathurai, and S. Rajendiran, “Novel segmentation and classification algorithm for detection of tomato leaf disease,” Concurrency and Computation: Practice and Experience, vol. 35, no. 12, 2023.

[4] S. H. Lee, H. Go¨eau, P. Bonnet, and A. Joly, “New perspectives on plant disease characterization based on deep learning,” Computers and Electronics in Agriculture, vol. 170, p. 105220, 2020.

[5] H. Nazki, S. Yoon, A. Fuentes, and D. S. Park, “Unsupervised image translation using adversarial networks for improved plant disease recognition,” Computers and Electronics in Agriculture, vol. 168, p. 105117, 2020.

[6] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Computers and Electronics in Agriculture, vol. 145, p. 311–318, 2018.

[7] M. Agarwal, S. K. Gupta, and K. Biswas, “Development of efficient cnn model for tomato crop disease identification,” Sustainable Computing: Informatics and Systems, vol. 28, p. 100407, 2020.

[8] P. Wspanialy and M. Moussa, “A detection and severity estimation system for generic diseases of tomato greenhouse plants,” Computers and Electronics in Agriculture, vol. 178, p. 105701, 2020.

[9] Pawar, U.B., Bhirud, S.G., Kolhe, S.R. (2020). Light Scattering Study on Protocols and Simulators Used in Automotive Application(s). In: Iyer, B., Deshpande, P., Sharma, S., Shiurkar, U. (eds) Computing in Engineering and Technology. Advances in Intelligent Systems and Computing, vol 1025. Springer, Singapore. https://doi.org/10.1007/978-981-32-9515-5_16

[10] A. Ali, “Plantvillage dataset,” Sep 2019