Tomato Leaf Disease Detection

Ms. Roopa[1], Ayush C[2]

[1] Assistant Professor, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru

[2]U.G. Student, Computer Science Department, JSS College of Arts Commerce and Science Ooty-Road, Mysuru

---***---

Abstract - Tomato plants are particularly vulnerable to leaf diseases, which can take a serious toll on crop yield and quality if not caught early. Traditionally, farmers and agricultural experts rely on manual inspection to spot these diseases a method that can be both slow and prone to mistakes. To streamline this process, our project introduces an automated system that uses machine learning and image processing to detect tomato leaf diseases more accurately and efficiently.

We worked with a dataset of tomato leaf images that includes both healthy leaves and those affected by diseases like Early Blight, Late Blight, Leaf Mold, and Target Spot. Using this data, we trained a Convolutional Neural Network (CNN) to classify the images with high accuracy. To improve the model's learning, we applied standard preprocessing steps such as image augmentation, resizing, and normalization. Our experiments show that the system performs well, offering reliable early detection of various tomato leaf diseases. By providing timely insights, this tool can help farmers take preventive action sooner, potentially reducing crop loss and improving productivity. In the future, we hope to expand this work into mobile platforms to make real-time disease detection more accessible in the field.

Key Words: Tomato leaf diseases, disease detection, deep learning, CNN, image processing, Late blight, Leaf mold.

1. INTRODUCTION

Tomatoes are one of the most widely grown and consumed crops in the world, playing a major role in global food security and local economies. Not only are they a staple ingredient in countless dishes around the globe, but they also offer essential nutrients like vitamins A, C, and potassium, making them an important part of a healthy diet for millions of people.

Beyond their nutritional value, tomatoes are economically significant. Their cultivation provides income for farmers, supports jobs in agriculture and food processing, and con- tributes to the agricultural sector's growth in many countries. Given their global importance, maintaining the health and productivity of tomato crops is essential for both economic sustainability and food availability.

However, tomato plants are vulnerable to a wide range of diseases, many of which affect the leaves and can severely reduce both crop yield and quality. Fungal, bacterial, and viral infections—such as early blight, late blight, Septoria leaf spot, tomato mosaic virus, and tomato yellow leaf curl virus—can stunt plant growth, damage leaves and fruits, or even destroy entire crops. These diseases not only impact food supply but also create financial strain, especially for small-scale farmers who depend heavily on successful harvests.

2. LITERATURE SURVEY

In recent years, a significant amount of research has been conducted in the field of plant disease detection, employing both traditional machine learning and modern deep learning approaches.

In 2019, Amrita S. Tulshan and Nataasha Raul proposed a method using the K-Nearest Neighbor (KNN) algorithm and achieved an impressive accuracy of 98.56% for identifying plant leaf diseases. Going back to 2013, Arti N. Rathod, Bhavesh Tanawal, and Vatsal Shah explored various image processing techniques for detecting leaf diseases, laying foun- dational work in this area.

In 2018, Prajwala TM and colleagues introduced a mod- ified version of the LeNet Convolutional Neural Network (CNN) to detect diseases in tomato leaves, reporting an accuracy of around 94–95%. Furthering this research, Surampalli Ashok and his team in 2020 applied multiple models—AlexNet, Artificial Neural Networks (ANN), and CNN—for tomato leaf disease detection. They achieved ac- curacies of 95.75%, 92.94%, and 98.12% respectively. Similarly, Mohit Agarwal and his team used a CNN model in their 2020 study and obtained an accuracy of 91.2%.Halil Durmus¸ and colleagues evaluated AlexNet and SqueezeNet models for tomato disease classification, reporting accuracies of 95.65% and 94.3% respectively . In another notable study from 2018, Konstantinos P. Ferentinos used a VGG-based model and achieved an exceptional accuracy of 99.48%. On the other hand, Alvaro Fuentes and his team developed a real-time tomato disease and pest recognition sys- tem using VGG16, which resulted in an accuracy of 83.06% . In 2019, Geetharamani and Arun Pandian proposed a deep CNN model with nine layers for identifying plant leaf diseases and achieved 96.46% accuracy. That same year, Peng Jiang and colleagues introduced an improved CNN approach for detecting apple leaf diseases, utilizing VGG-FCN-VD16 and VGG-FCN-S architectures, which yielded accuracies of 97.95% and 95.12% respectively [?]. Likewise, Xihai Zhang and his team worked on maize leaf disease detection using a GoogleNet-based model and achieved 98.9% accuracy .

Melike Sardogan and her team combined CNN with the Learning Vector Quantization (LVQ) algorithm in 2018, re- sulting in an average accuracy of 86% [?]. Meanwhile, Yang Lu and collaborators focused on rice disease identification in 2017 using deep CNNs, reaching an accuracy of 95.48%. Jiang Lu and his team developed an in-field wheat disease diagnosis system using VGG-FCN architectures, obtaining recognition accuracies of 97.95% and 95.12% during five-fold cross-validation. Lastly, in 2020, Utkarsha N. Fulari and colleagues used CNN models on grape and strawberry leaf datasets and achieved outstanding accuracies of 99.7% and 100% respec- tively, highlighting the remarkable potential of deep learning in crop disease detection .

3. PROPOSED METHODOLOGY

The methodology adopted in this study consists of several important stages to ensure accurate and efficient detection of tomato leaf diseases using deep learning techniques.

The first step is to obtain a collection of tomato leaf photos from the Kaggle platform. This collection contains photos of healthy and damaged tomato leaves from various categories.

. Since the raw data cannot be fed directly into the models, it undergoes a preprocessing stage. During preprocessing, all images are resized—150×150 pixels for the Convolutional Neural Network (CNN) model and 224×224 pixels for the ResNet-50 model. In addition, data augmentation techniques such as flipping, rotation, zooming, and shifting are applied to enrich the dataset and help prevent overfitting.

Once the data is preprocessed, it is divided into training, validation, and testing subsets. This ensures that the model can learn effectively from the training data, validate its per- formance during training, and finally be evaluated on unseen data.

Next, a CNN model is constructed, consisting of multiple convolutional and pooling layers followed by fully connected layers. In parallel, a ResNet-50 architecture is also imple- mented using transfer learning to compare results. Both models are trained on the dataset for multiple epochs, and their performance is monitored using accuracy and loss metrics.

Finally, the trained models are evaluated using the testing data, and the results are analyzed based on overall accuracy and their ability to correctly classify various types of tomato leaf diseases.

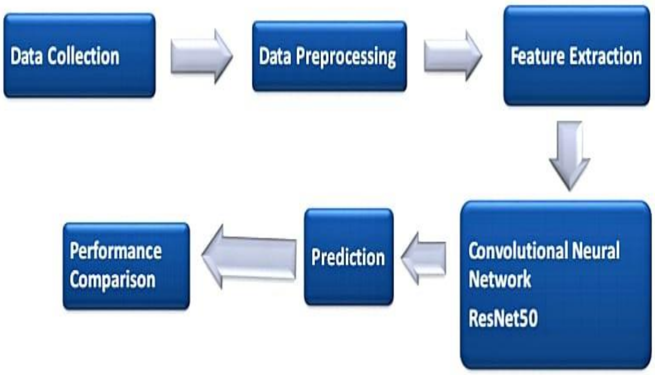


Fig.3.1: Architecture diagram of the proposed system.

4. SYSTEM DESIGN

The process of detecting tomato leaf diseases begins with data collection. For this study, a dataset of tomato leaf images was sourced from Kaggle, which provided a wide variety of healthy and diseased leaf samples in raw image format. After collecting the data, the next step was preprocessing.Once the data was collected, the next step was preprocess- ing. This included resizing the images to ensure uniform input dimensions for the neural networks. Specifically, images were resized to 150×150 pixels for the CNN model and to 224×224 pixels for the ResNet50 model to match their respective input layer requirements.

To enhance model performance and reduce overfitting, data augmentation techniques were applied. This involved generating additional training samples through operations like rotation, flipping, zooming, and shifting, which helped the model generalize better to unseen data.

Following preprocessing, relevant features were automatically extracted using the network layers. Finally, the dataset was split into training and testing sets to evaluate the performance of the models. These steps laid the groundwork for building a robust and accurate

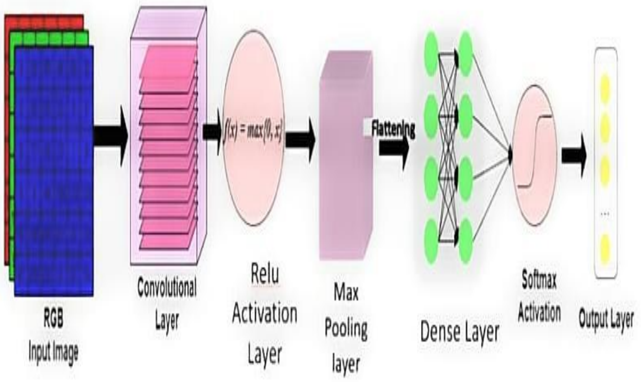


Fig.4.1: System Architecture Diagram

5. Experiments and Results

To evaluate the effectiveness of our tomato leaf disease detection model, a series of experiments were carried out using two architectures: a custom Convolutional Neural Network (CNN) and a pre-trained ResNet50 model. The models were trained and tested using the Kaggle tomato leaf dataset, which contains over 16,000 images across ten categories, including healthy and diseased leaves.

During training, multiple experiments were conducted with different epoch values (10, 20, and 50) to observe how performance evolved over time. For the CNN model, accuracy improved significantly with more training epochs — achieving 64% accuracy at 10 epochs, 94% at 20 epochs, and 97% at 50 epochs. These results suggest that the CNN model effectively learns complex features from the tomato leaf images as training progresses.

In comparison, the ResNet50 model, which utilizes transfer learning, achieved high accuracy even with fewer training epochs due to its deep architecture and pre-learned weights. The use of data augmentation and proper preprocessing tech- niques played a crucial role in boosting model performance and preventing overfitting.

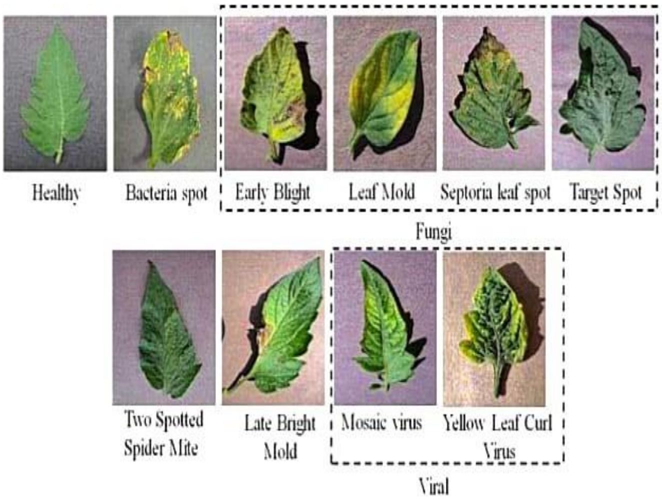
The results clearly demonstrate that deep learning models, particularly CNNs and ResNet50, are well-suited for detecting and classifying diseases in tomato leaves with high precision.

A. Visualization

This chapter presents the outcomes of our experimental analysis, highlighting both the training progress and the pre- dictive performance of our CNN-based model for tomato leaf disease detection. The results include visualizations of the model's training accuracy and loss over multiple epochs, offering insights into how well the model has learned to distinguish between healthy and diseased leaves. These findings validate the effectiveness of the proposed deep learning approach in identifying different types of tomato leaf diseases with a high degree of accuracy.

fig5.1visualization of tomato leaf images

B. Tomato Leaf Disease Detection



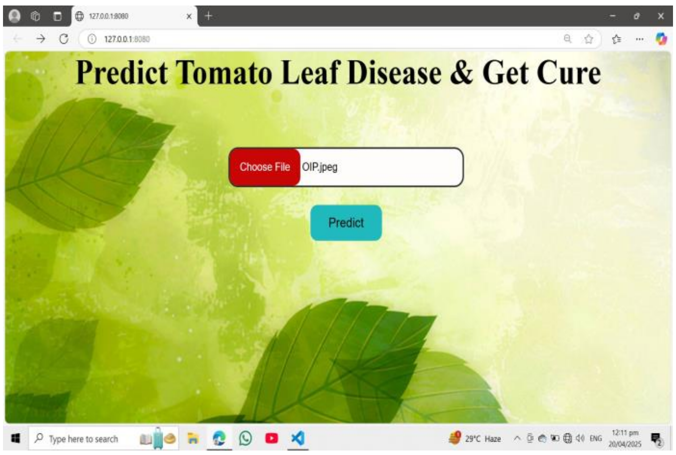


Fig5.2. Tomato Leaf Disease Detection

C. Tomato Leaf Disease Detection result

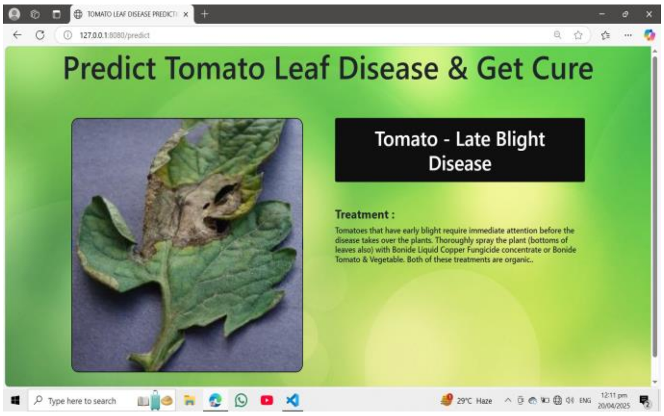


Fig 5.3 Tomato Leaf Disease Detection result

6. CONCLUSION

This study aimed to develop an automated system for detecting diseases in tomato leaves using deep learning, particularly Convolutional Neural Networks (CNNs). The model was trained on the Plant Village dataset, which included labeled images of both healthy and diseased tomato leaves across ten distinct categories.

Training the CNN model at 10, 20, and 50 epochs demonstrated a consistent improvement in accuracy with increased training. The best performance was observed at 50 epochs, where the model achieved an accuracy of 97%, highlighting the effectiveness of deep learning for this classification task.

Preprocessing techniques, such as resizing and data augmentation, were crucial in enhancing model performance and preventing overfitting. These steps exposed the network to a diverse range of inputs, enabling better generalization on unseen data.

Overall, the proposed CNN-based method shows great promise for practical applications in agriculture. With high classification accuracy, it could serve as a valuable tool for early disease detection, assisting farmers in taking timely actions to protect their crops and minimize yield loss.

For future development, the system could be enhanced by incorporating real-time image capture, extending the dataset with field-acquired images, and experimenting with more advanced CNN architectures like ResNet or EfficientNet. Additionally, developing a mobile or web-based application could make this technology more accessible to end-users, promoting smarter and more efficient farming practices.

REFERENCES

[1] Tulshan, Amrita S., and Nataasha Raul. ”Plant leaf disease detection using machine learning.” 2019 10th In- trinational Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2019.

[2] Rathod, Arti N., Bhavesh Tanawal, and Vatsal Shah. ”Image processing techniques for detection of leaf disease.” International Journal of Advanced Research in Computer Science and Software Engineering 3.11 (2013).

[3] Tm, Prajwala, et al. ”Tomato leaf disease detection using convolutional neural networks.” 2018 eleventh international conference on contemporary computing (IC3). IEEE, 2018.

[4] Ashok, Surampalli, et al. ”Tomato Leaf Disease Detec- tion Using Deep Learning Techniques.” 2020 5th International Conference on Communication and Electronics Systems (IC- CES). IEEE, 2020.

[5] Agarwal, Mohit, et al. ”ToLeD: Tomato leaf disease de- tection using convolution neural network.” Procedia Computer Science 167 (2020): 293-301.

[6] Durmus¸, Halil, Ece Olcay Gu¨nes¸, and Mu¨rvet Kırcı. ”Disease detection on the leaves of the tomato plants by using deep learning.” 2017 6th International Conference on Agro- Geoinformatics. IEEE, 2017.

[7] Ferentinos, Konstantinos P. ”Deep learning models for plant disease detection and diagnosis.” Computers and elec- tronics in agriculture 145 (2018): 311-318.

[8] Fuentes, Alvaro, et al. ”A robust deep-learning-based detector for real-time tomato plant diseases and pests recog- nition.” Sensors 17.9 (2017): 2022.

[9] Geetharamani, G., and Arun Pandian. ”Identification of plant leaf diseases using a nine-layer deep convolutional neural network.” Computers Electrical Engineering 76 (2019): 323-

[10] Jiang, Peng, et al. ”Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks.” IEEE Access 7 (2019): 59069-
59080.

[11] Zhang, Xihai, et al. ”Identification of maize leaf diseases using improved deep convolutional neural networks.” Ieee Access 6 (2018): 30370-30377.