Detection of Tomato Leaf Infections

Ms. Roopa[1], Darshan S[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 - Plant diseases pose a significant threat to agriculture by reducing crop yield and quality, often without early signs noticeable to the naked eye. Among these, tomato crops are especially vulnerable to a variety of diseases that can quickly spread and impact production. This study proposes an effective and automated approach to detecting tomato leaf diseases using deep learning, specifically Convolutional Neural Networks (CNNs).

The system is designed to classify tomato leaf images into ten categories: healthy, yellow leaf curl virus (YLCV), bacterial spot (BS), early blight (EB), leaf mold (LM), septoria leaf spot (SLS), target spot (TS), two-spotted spider mite spot (TSSMS), mosaic virus (MV), and late blight (LB). A dataset comprising 16,021 labeled tomato leaf images was used for training. The model was trained across three configurations: 10, 20, and 50 epochs, yielding accuracy rates of 64%, 94%, and 97% respectively.

The results show a clear improvement in accuracy with extended training, demonstrating the potential of CNN-based models for practical plant disease detection. This technology can play a vital role in assisting farmers with early disease identification, ultimately leading to better crop management and higher agricultural productivity.

***Key Words*:** Convolutional Neural Networks, Max Pooling layer, Yellow Leaf Curl Virus, Bacterial Spot, Early Blight, Leaf Mold, Spectorial Leaf Spot, Target Spot, Two spotted Spider Mite Spot, Mosaic Virus, Late Blight

1. INTRODUCTION

Tomatoes are one of the most widely cultivated and consumed crops globally, contributing significantly to both the food industry and agricultural economy. However, tomato plants are highly susceptible to various diseases, which can severely impact crop yield and quality, resulting in major economic losses for farmers. Among these diseases, grey leaf spot is particularly harmful—it damages the leaves, disrupts the plant's photosynthesis process, and ultimately reduces fruit production. This disease progresses through four key stages: contact, invasion, latency, and onset. Early detection is essential to prevent its spread and minimize damage.

Traditional methods of disease detection often involve man- ual inspection, which is not only time-consuming and labor- intensive but also impractical for large-scale farming. These methods may not consistently detect diseases in their early stages, especially without expert knowledge. To overcome these limitations, image processing and deep learning tech- niques have emerged as effective alternatives. They provide faster, more accurate, and scalable solutions for identifying plant diseases.

In this study, we propose a deep learning-based approach using Convolutional Neural Networks (CNNs) to automatically detect and classify tomato leaf diseases. The model is trained on a dataset of 16,021 images representing ten categories of healthy and diseased tomato leaves. Each image is resized to 256×256 pixels and split into training, testing, and validation sets. This approach leverages CNNs' ability to extract mean- ingful patterns from visual data, enabling efficient and accurate disease classification. The implementation of this technology can support early intervention, reduce crop loss, and improve productivity for tomato growers.

2. LITERATURE SURVEY

The detection of plant diseases has become an increasingly important area of study, especially with the rise of machine learning and deep learning techniques. Traditional computer vision methods have long been used to extract features such as color, texture, and shape from plant images. However, these approaches often rely on expert domain knowledge and may lack scalability and accuracy when applied to large datasets or real-time scenarios.With the rapid growth of artificial intelligence, many re- searchers have turned to deep learning models, especially Convolutional Neural Networks (CNNs), for disease detection in plants. These models learn patterns directly from images, eliminating the need for hand-crafted feature extraction and significantly improving performance.

In 2019, Amrita S. Tulshan and Nataasha Raul applied the K-Nearest Neighbor (KNN) algorithm for plant disease detec- tion, achieving an impressive accuracy of 98.56%. Earlier, in 2013, Arti N. Rathod et al. explored various traditional image processing techniques for identifying infected leaves.

Several studies have focused specifically on tomato leaf disease detection. In 2018, Prajwala TM and colleagues used a modified LeNet CNN model and achieved an average accu- racy of 94–95%. In 2020, Surampalli Ashok et al. evaluated AlexNet, Artificial Neural Networks (ANN), and CNN models for tomato diseases, reporting accuracies of 95.75%, 92.94%, and 98.12%, respectively. Another 2020 study by Mohit Agar- wal and his team used CNNs to detect tomato leaf diseases, reaching an accuracy of 91.2%.

Halil Durmus¸ et al. tested AlexNet and SqueezeNet models on tomato leaf images and achieved 95.65% and 94.3% accuracy, respectively. Konstantinos Ferentinos applied the VGG architecture to various plant diseases and obtained a high accuracy of 99.48%. Alvaro Fuentes and his colleagues used VGG-16 for real-time detection of tomato plant diseases and pests, achieving 83.06% accuracy.

In 2019, Geetharamani and Arun Pandian developed a nine- layer deep CNN to detect leaf spot diseases and achieved 96.46% accuracy. PENG JIANG et al. used improved VGG- FCN-VD16 and VGG-FCN-S models for real-time apple leaf disease detection and achieved recognition accuracies of 97.95% and 95.12%, respectively. Similarly, XIHAI ZHANG and colleagues applied GoogLeNet for maize leaf disease detection.

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% respectively, highlighting the remarkable potential of deep learning in crop disease detection .

3. PROPOSED METHODOLOGY

This section outlines the methodology adopted for detecting tomato leaf diseases using Convolutional Neural Networks (CNN). The entire process is divided into several key stages: data collection, preprocessing, model design, training, evaluation, and prediction. Each of these steps is described in detail below.

A. Data Collection

The dataset was sourced from publicly available image repositories that contain labeled images of tomato leaves affected by various diseases, as well as healthy ones. Each image is annotated with its corresponding disease class, such as Early Blight, Late Blight, or Healthy.

B. Data Preprocessing

Before training the model, the images underwent preprocessing to ensure consistency and improve model performance. This included:

- Resizing all images to a fixed resolution (e.g., 224x224 pixels).
- Normalizing pixel values to the range [0, 1].
- Applying data augmentation techniques like rotation, flipping, and zooming to improve generalization.

C. Model Architecture

A Convolutional Neural Network (CNN) was designed and trained to classify tomato leaf images into their respective categories. The architecture typically includes:

- Multiple convolutional layers for feature extraction.
- Max-pooling layers to reduce dimensionality.

- Dropout layers to prevent overfitting.
- Dense (fully connected) layers for final classification.

D. Training and Validation

The dataset was split into training, validation, and test sets (e.g., 70% for training, 15% for validation, and 15% for testing). The training process involved:

- Compiling the model with a suitable optimizer (e.g., Adam) and loss function (e.g., categorical crossentropy).
- Training the model for a predefined number of epochs.
- Monitoring accuracy and loss on both training and vali- dation sets to avoid overfitting.

E. Evaluation

After training, the model was evaluated on the test set to assess its performance. Metrics such as accuracy, precision, recall, and F1-score were used to analyze the results.

F. Prediction

Finally, the trained CNN model was used to predict diseases from new, unseen images of tomato leaves. The model outputs the most probable disease class for a given input image

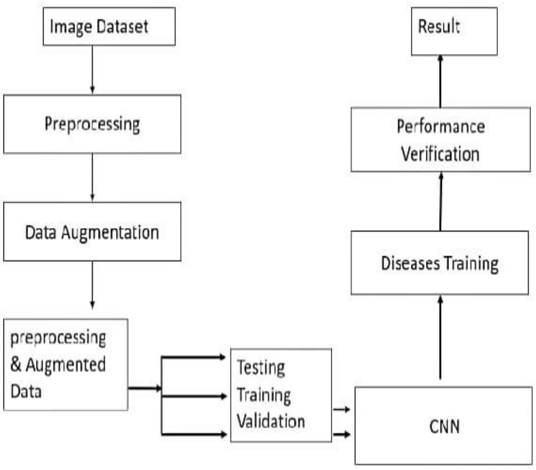

Fig 3.1 Architecture diagram of the proposed system

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

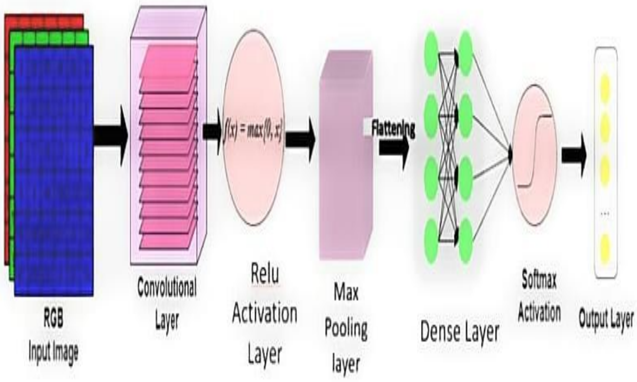

Fig.4.1: System Architecture Diagram

5 . 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

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.

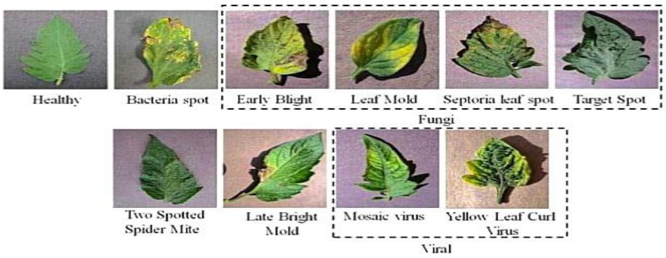

Fig 5.1 Visual

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.

B. Tomato Leaf Disease Detection

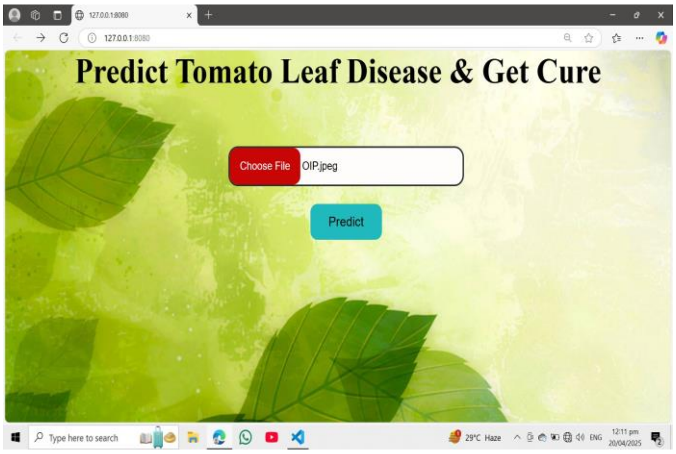

Fig5.2. Tomato Leaf Disease Detection

C. Tomato Leaf Disease Detection result

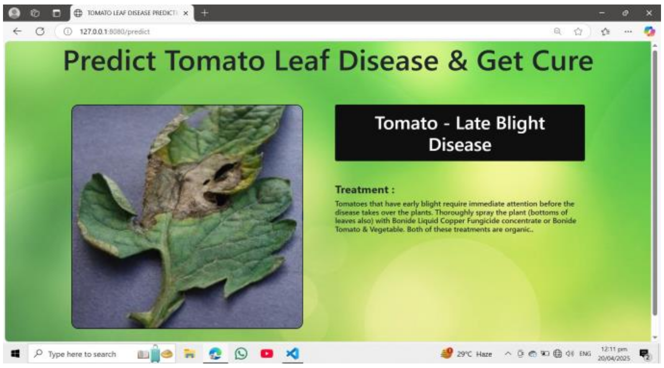

Fig 5.3 Tomato Leaf Disease Detection result

6.CONCLUSION

In this project, we developed a deep learning-based system for the detection of tomato leaf diseases using Convolutional Neural Networks (CNNs). The model was trained on a diverse dataset of tomato leaf images and was able to successfully identify ten common tomato plant diseases with high accuracy. The system is designed to be simple, fast, and accessible, offering real-time predictions through a web-based interface. This makes it particularly useful for farmers and agricultural experts who need quick and reliable disease identification in the field.By automating the disease detection process, this approach has the potential to reduce crop loss, improve yield quality, and support more effective plant health monitoring. Overall, the project demonstrates how artificial intelligence and deep learning can play a significant role in modern agriculture.

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.