Early Disease Detection & Classification of Tomato Plant

Sagar Agarwal[1]	Dr J P Patra[2]
Department of Computer Science & Engineering, Shri Shankaracharya Institute of Professional Management & Technology, Raipur, India sagar.agarwal@ssipmt.com	Professor, Computer Science & Engineering, Shri Shankaracharya Institute of Professional Management & Technology, Raipur, India patra.jyotiprakash@gmail.com

*Corresponding Author: **patra.jyotiprakash@gmail.com**

Article History: Received: 05.10.2022 Revised: 23.12.2022 Accepted: 17.1.2023

Abstract

Tomato is a highly valuable crop in the Indian market, cultivated in large quantities. However, tomato plants are susceptible to various diseases, which can lead to significant reductions or complete destruction of the crop, resulting in substantial economic losses for farmers. Therefore, the early and accurate detection of these diseases is crucial to minimize such losses. In recent years, numerous methods have been proposed to detect plant diseases, leveraging advancements in technology. Despite these efforts, achieving high accuracy in detecting tomato plant diseases remains a challenge. In this paper, we present a comprehensive study where we explore the effectiveness of deep learning models in detecting and classifying diseases in tomato plants. Specifically, we investigate four different pre-trained deep learning models, namely VGG-19, ResNet-50, Inception V3, and InceptionResNetV2. These models are trained using three different optimizers and various learning rates. By leveraging these models, we aim to develop an efficient framework that can accurately identify and classify diseases in tomato plants, thereby reducing production and economic losses. The proposed framework holds significant potential in revolutionizing the early detection and classification of tomato plant diseases. The utilization of deep learning models and optimization techniques contributes to improving the accuracy and efficiency of disease detection systems. This research aims to provide valuable insights for farmers and stakeholders in the agricultural industry, enabling them to take timely and appropriate actions to mitigate the impact of diseases on tomato crop yields.

Keywords: *Early disease detection, disease classification, tomato plants, deep learning models, pre-trained models, optimization techniques, agricultural industry.*

1. INTRODUCTION

Agriculture plays a vital role in the Indian economy, with a significant portion of the population depending on it for their livelihood. Among the various crops cultivated in India, tomatoes hold immense importance as they are widely consumed and produced across the country. In fact, India ranks as the second-largest producer of tomatoes globally, with an annual production of 18,399,000 tonnes in 2021. In the state of Chhattisgarh, tomato production reached 1149 MT, securing the eighth position in terms of output. However, tomato farmers often encounter substantial economic losses due to various diseases that affect the crop. Early detection and timely treatment of these diseases are crucial in order to prevent significant losses and preserve the health of the plants. This project aims to address this issue by developing a disease detection system capable of identifying diseases in tomato plants based on the symptoms exhibited on their leaves. The methodology employed in this study focuses on the most common diseases affecting tomato plants, including Leaf Mold, Early Blight, Spider Mites (Two-Spotted Spider Mites), Tomato Healthy, and Septoria Leaf Spot. Accurate prediction of these diseases is essential as each disease requires a specific remedy, and the application of an incorrect treatment can have adverse effects on the plant. Human observation alone is often insufficient for accurately detecting these diseases, as it can be challenging to discern the symptoms with a single glance. This can result in incorrect assumptions about the disease, leading to the use of inappropriate treatments and, ultimately, the loss of the entire plant. Therefore, it is imperative to develop a robust model that can predict diseases in tomato plants with the utmost accuracy. The proposed model aims to leverage advanced techniques, such as deep learning, to accurately identify and classify diseases based on visual cues. By employing such a model, farmers will be equipped with a reliable tool that can aid in the early detection of diseases, enabling timely interventions and appropriate treatments. Ultimately, this research endeavors to provide a practical solution that empowers tomato farmers to mitigate losses and ensure the health and productivity of their crops.

2. LITERATURE SURVEY

Prajwala TM [1] proposed a model of LeNet which is a simple CNN model which consists of convolutional, activation, pooling, and connected layers. The accuracy of the model is 94% which needs to be further increased and also we can inform them about the solution to these diseases while detecting them.

As discussed in [2] they proposed a comparison between 3 pre-trained models SVM & k-nearest with each model tested with 2 feature extractions - Inception V3, VGG - 16, VGG- 19. This model gave an accuracy of 96.8 %. The accuracy can be improved and work can be easily extended to identify the diseases for other crops as well.

In this paper [3], three popular classifiers were assessed that is SVM, RF, and MLR. The highest accuracy between different classifiers obtained was 92.4% with MLR. The proposed model was able to detect 6 diseases it can be further developed to identify more diseases with more accuracy.

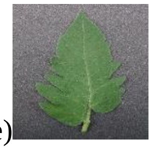
The result discussed in paper [4] confirmed that DNN deep learning classifier gives an improved accuracy of 86.18%. The accuracy can be easily improved.

As discussed in paper [5] they used a deep neural network model for detecting and classifying tomato plant leaf diseases into predefined categories and the proposed model was compared to VGG and ResNet versions. It achieved an accuracy rate of 98.43%. The dataset on which it was performed was less so we need to gather more data to test it over the same

3. METHODOLOGY

3.1 Dataset description

We have selected the images dataset from Kaggle – Plant Village Tomato Lead Dataset which consists of 14,529 images having 5 different classes. It has a) 1702 images of Tomato Septoria Leaf Spot b) 800 images of early blight disease c) 761 images of Tomato Leaf mold d) 1341 images of spider mites e) 1273 healthy tomato leaf images. The size of the images which we used was 256 x 256 and all of the images were of RGB color space.

a)  b)  c)  d)  e) 

3.2 Dataset Acquisition

To train and evaluate our disease detection and classification model, a comprehensive dataset of tomato plant images is required. We collect a diverse set of high-resolution images representing healthy tomato plants and various diseased conditions. The dataset should include images exhibiting symptoms of Leaf Mold, Early Blight, Spider Mites (Two-Spotted Spider Mites), Tomato Healthy, and Septoria Leaf Spot. Additionally, the dataset should be properly annotated, indicating the disease labels for each image.

3.3 Preprocessing

Preprocessing of the acquired dataset is crucial to ensure optimal model performance. Initially, we perform data cleaning, removing any corrupt or irrelevant images. Next, we resize the images to a standardized resolution to ensure consistency across the dataset. Additionally, we employ techniques such as normalization and augmentation to enhance the robustness and generalization of the model.

3.4 Model Selection

In this study, we focus on deep-learning models for disease detection and classification in tomato plants. We consider several state-of-the-art pre-trained models, including VGG-19, ResNet-50, Inception V3, and InceptionResNetV2. These models are known for their ability to extract meaningful features from images and exhibit high performance in computer vision tasks. We compare the performance of these models to identify the most suitable one for our specific task.

3.5 Transfer Learning

To leverage the pre-trained models effectively, we employ transfer learning. We fine-tune the selected pre-trained model on our tomato plant dataset. During the fine-tuning process, we freeze the early layers of the network, which are responsible for capturing generic features, and fine-tune the later layers to adapt to the specific features related to tomato plant diseases. This approach enables efficient training even with limited annotated data.

3.6 Model Training

We split the preprocessed dataset into training, validation, and testing sets. The training set is used to train the model on the annotated images, while the validation set is utilized to optimize the model hyper parameters and prevent overfitting. We employ appropriate loss functions, such as categorical cross-entropy, and employ optimization algorithms, such as SGD, Adam & RMSprop to train the model. We experiment with different learning rates (0.01, 0.001, 0.004, and 0.008) and select the optimal values based on the validation performance.

3.7 Model Evaluation

Once the model training is complete, we evaluate its performance on the testing set, which contains unseen images. We measure various evaluation metrics, such as accuracy, precision, recall, and F1 score, to assess the model's disease detection and classification capabilities. We compare the results across different diseases and models to identify the strengths and limitations of the proposed approach.

To evaluate the performance of the model a set of quantitative metrics comprising precision, accuracy, recall & F1 score is used.

Accuracy = Number of correct predictions / Total number of predictions

Accuracy = (TP + TN) / (TP + FP + TN + FN)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F1 = 2 * (Precision * Recall) / (Precision + Recall)

Where TP = True Positives, TN = True Negatives, FP = False Positives & FN = False Negatives

3.8 Performance Comparison

To validate the effectiveness of our proposed approach, we compare the performance of our model with existing methods or alternative techniques for tomato plant disease detection. We analyze the accuracy, computational efficiency, and potential for real-time deployment to highlight the advantages of our methodology.

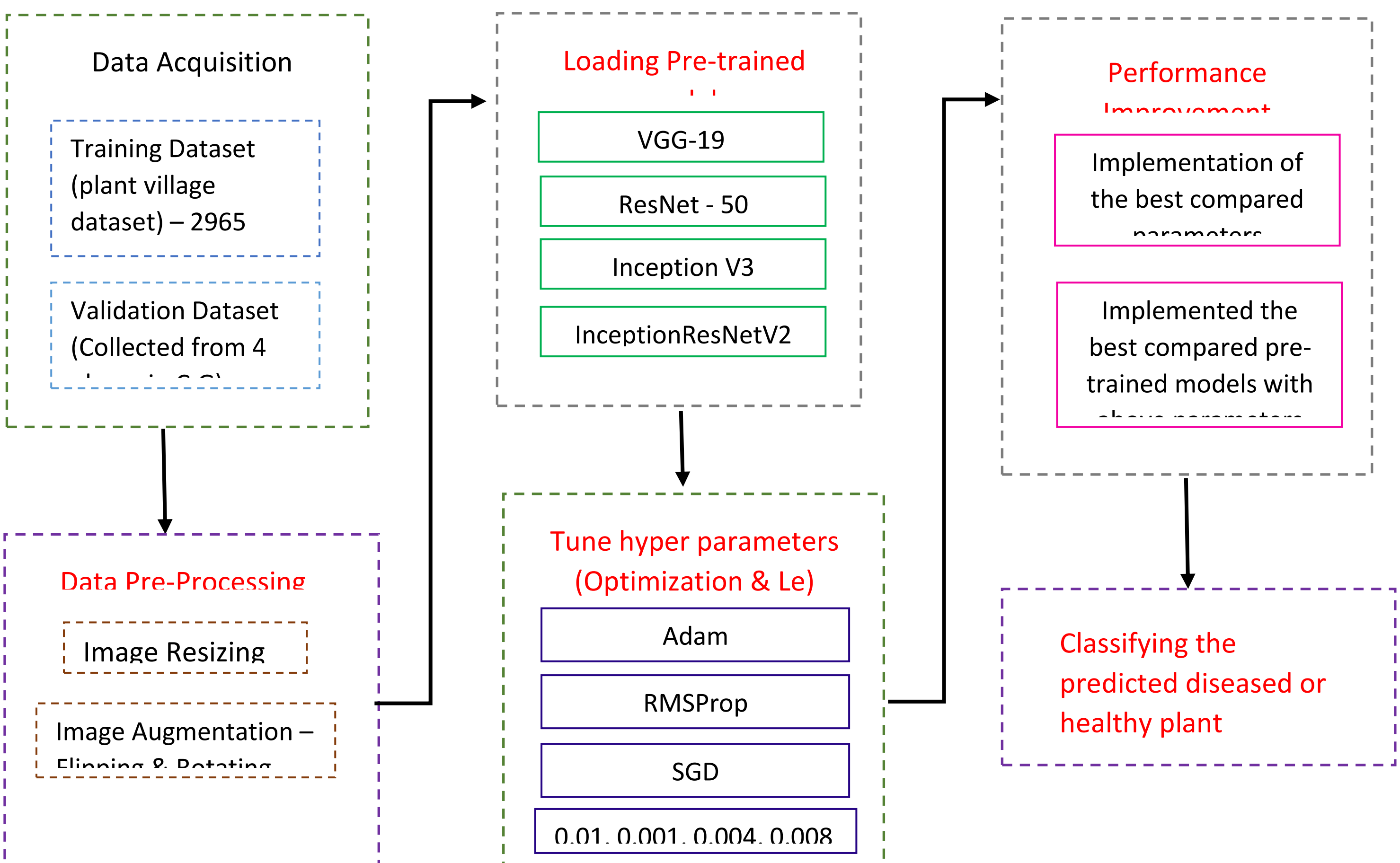


Figure 1: Proposed Model

4.IMPLEMENTATION

We used Flask to implement all four models, a framework that made it simple to incorporate the models into our user-friendly online application. This made it easy for users to access and make use of our application's potent tomato disease detection features.

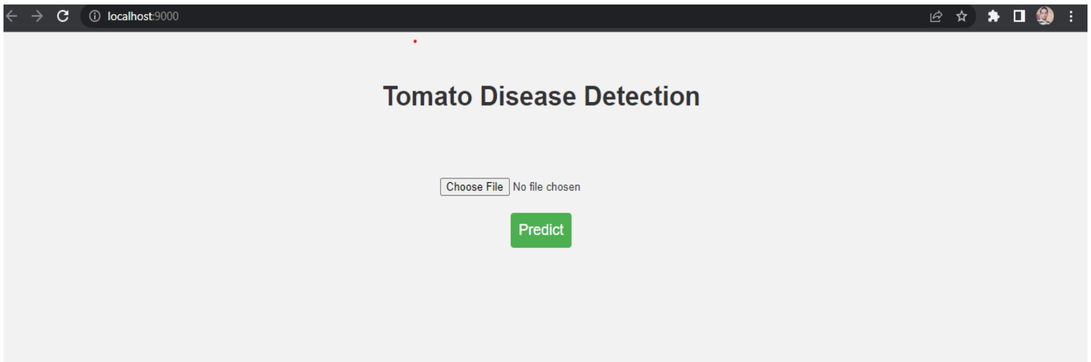


Figure 2: Webpage with implemented model

5.RESULT & CONCLUSION

Upon analyzing the results, it is evident that Adam optimizer consistently outperformed other optimizers across all learning rates. Notably, at a learning rate of 0.001, the Adam optimizer achieved the highest accuracy compared to other optimizers, demonstrating its effectiveness in training the VGG-19 model for tomato disease detection as shown in the graph.

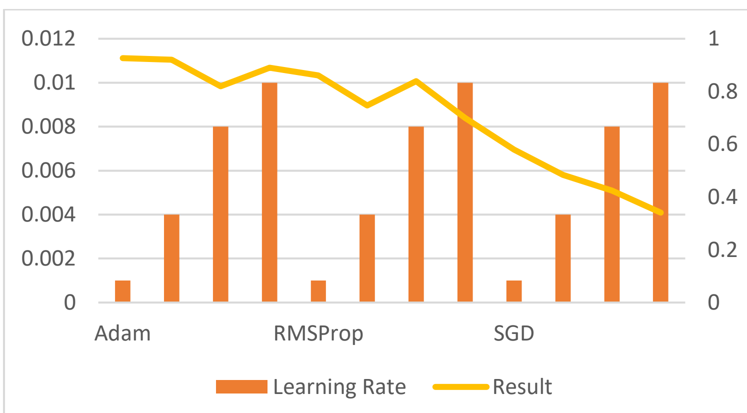
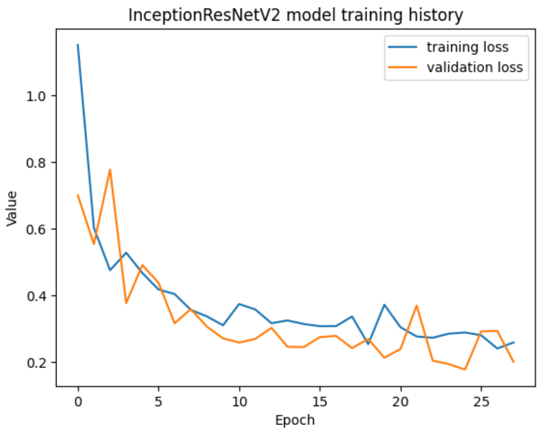
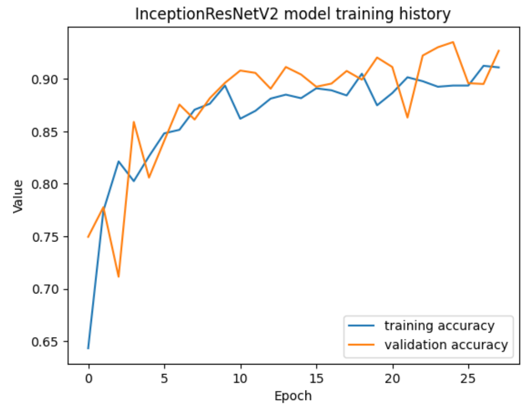


Figure 3: Chart for Optimizers vs Learning Rate on VGG-19 model





The performance of our proposed disease detection and classification system was evaluated using F1 scores, which measure the model's accuracy in identifying and classifying diseases in tomato plants. The following F1 scores were obtained for each model:

Models	F1 score
VGG-19	0.3025
ResNet 50	0.2061
Inception V3	0.4201
InceptionResNet V2	0.5502

These F1 scores indicate the overall effectiveness of each model in correctly identifying and classifying the diseases present in tomato plants. Among the models evaluated, InceptionResNet V2 achieved the highest F1 score of 0.5502, demonstrating its superior performance in disease detection and classification tasks. Inception V3 also exhibited favorable results with an F1 score of 0.4061. Furthermore, our experiments revealed that using the ADAM optimizer with a learning rate of 0.001 yielded the best overall performance across the evaluated models. This configuration facilitated better convergence and optimization of the deep learning models, resulting in improved disease detection accuracy.

References

[1] P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi and S. G. Koolagudi, "Tomato Leaf Disease Detection Using Convolutional Neural Networks," *2018 Eleventh International Conference on Contemporary Computing (IC3)*, 2018, pp. 1-5, doi: 10.1109/IC3.2018.8530532.

[2] Tm, P., Pranathi, A., Saiashritha, K., Chittaragi, N.B., and Koolagudi, S.G. (2018) Tomato Leaf Disease Detection Using Convolutional Neural Networks. The 11th International Conference on Contemporary Computing, Noida, 2-4 August 2018, 1-5.

[3] Sreelatha, Dr. et al. “Managing the Tomato Leaf Disease Detection Accuracy Using Computer Vision-Based Deep Neural Network.” (2021).

[4] Chowdhury, Muhammad & Rahman, Tawsifur & Khandakar, Amith & Ibtehaz, Nabil & Khan, Aftab & Khan, Muhammad Salman & Al-Emadi, Nasser & Reaz, Mamun Bin Ibne & Islam, Mohammad & Ali, Sawal. (2021). Tomato Leaf Diseases Detection Using Deep Learning Technique. 10.5772/intechopen.97319.

[5] S. U. Habiba and M. K. Islam, "Tomato Plant Diseases Classification Using Deep Learning Based Classifier From Leaves Images," *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 2021, pp. 82-86, doi: 10.1109/ICICT4SD50815.2021.9396883.