Deep Learning for Early Detection of Tomato Leaf Diseases: A ResNet-18 Approach for Sustainable Agriculture

Asha M S¹, Yogish H K²

Department of Computer Science and Engineering, Christ University (Deemed to be University), Karnataka, India¹
Department of Information Science and Engineering, M S Ramaiah Institute of Technology, Karnataka, India²
Visvesvaraya Technological University, Belagavi, Karnataka, India-590018¹,²

Abstract—The paper explores the application of Convolutional Neural Networks (CNNs), specifically ResNet-18, in revolutionizing the identification of diseases in tomato crops. Facing threats from pathogens like Phytophthora infestans, timely disease detection is crucial for mitigating economic losses and ensuring food security. Traditionally, manual inspection and labour-intensive tests posed limitations, prompting a shift to CNNs for more efficient solutions. The study uses a well-organized dataset, employing data preprocessing techniques and ResNet-18 architecture. The model achieves remarkable results, with a 91% F1 score, indicating its proficiency in distinguishing healthy and unhealthy tomato leaves. Metrics such as accuracy, sensitivity, specificity, and a high AUC score on the ROC curve underscore the model’s exceptional performance. The significance of this work lies in its practical applications for early disease detection in agriculture. The ResNet-18 model, with its high precision and specificity, presents a powerful tool for crop management, contributing to sustainable agriculture and global food security.

Keywords—Convolution neural networks; tomato crop health; deep learning; binary classification; disease detection

I. INTRODUCTION

Tomato (Solanum Lycopersicon) holds an important place in agriculture, food and cooking worldwide. Known for its bright red color and its many uses in dishes such as Mediterranean pastas, Asian curries, and American tomato sauces, the tomato has become an important part of world cuisines [1]. In addition to its gastronomic importance, tomato is an important agricultural product that makes a significant contribution to global food production [2].

Despite the diversity and importance of tomatoes, tomatoes face many disease threats, such as late blight caused by Phytophthora infestans and fungal diseases that cause molds. The impact of these diseases on the tomato crop poses a constant risk to agriculture and can lead to severe economic and food shortages [3] [4]. Timely and accurate identification of these diseases is important for effective control and prevention [5] [6].

The method of detecting tomato diseases has always relied on manual inspection and laborious experiments that have their limitations. Depending on the interpreter, visual inspection may not be necessary for early detection of disease [7] [8] [9] [10] [11] [12]. While the tests are accurate, they are time-consuming and expensive, making it difficult to meet the urgent needs of today’s agriculture. To solve these problems, artificial intelligence (AI) has been transformed into agriculture in recent years, especially thanks to advances in deep learning.

Deep learning is a category of machine learning that uses multiple layers of artificial neural networks to solve complex problems. In deep learning, convolutional neural networks (CNN) have become powerful tools for data visualization and have become relevant in many fields [13] [14] [15] [16]. CNNs are characterized using convolutional techniques and are good at extracting relevant features from images, making them ideal for tasks such as image recognition and classification.

This change also extends to agriculture, where CNNs provide a way to quickly and accurately identify diseases, pests, and overall crop health. While this change applies to many crops, we focus on early detection of tomato leaf diseases. Against this problem, ResNet-18 architecture stands out as a light source that measures computational power and accuracy. Our research uses the ResNet-18 architecture to explore the potential of deep learning to transform the tomato plants system and even the permaculture system [17] [18] [19] [20] [21].

Our work set out on a journey to combine the fundamental world of tomatoes with the transformative power of deep learning. The stakes are high as we try to provide fast, effective solutions to ongoing challenges like growing healthy tomatoes, benefiting farmers, fields, and people around the world who depend on this versatile and important fruit. In the following sections, we will describe the process, results and conclusions of our research, which leads to a general discussion about permaculture and its important role in shaping the future of intelligence [22][23][24][25].

II. LITERATURE SURVEY

These days, there is a lot of research being done in the broad field of image processing applications for plant disease detection and classification. For the prompt identification of plant diseases, these applications are helpful. For any plant, diseases like fungus, bacteria, and viruses can be fatal. In Sabrol et al.’s [26] study, tomato late blight, Septoria spot, bacterial spot, bacterial canker, tomato leaf curl, and healthy tomato plant leaf and stem images are categorized into five categories. The categorization process involves removing
characteristics related to color, shape, and texture from images of healthy and unhealthy tomato plants. Following the segmentation step comes the feature extraction. It is extracted and put into the classification tree from segmented images.

A three-compact convolutional neural network (CNN) pipeline for the automatic detection of tomato leaf diseases has been proposed by Attallah et al. [28]. In order to obtain a more streamlined and sophisticated representation, deep features are extracted from the last fully connected layer of the CNNs using transfer learning. Then, in order to take use of each CNN structure, it combines features from the three CNNs. It then selects and creates an extensive feature set of smaller dimensions using a hybrid feature selection technique. The process for identifying tomato leaf diseases uses six classifiers. In order to confirm the suggested pipeline's ability to compete, the experimental results are also compared with earlier studies on the classification of tomato leaf diseases.

The circumstances of a tomato plant have been determined using a basic CNN model that contains eight hidden layers. When compared to other traditional models, the suggested strategies produce optimal results [24] [27] [29] [30] [31]. The image processing system recognizes and categorizes tomato plant illnesses using deep learning techniques. Here, the author implemented a full system using CNN and the segmentation technique. To attain higher accuracy, a CNN model modification has been implemented.

## III. Methods

### A. Datasets

The foundation of our classification model lies in the dataset of tomato leaf images, which we've organized into two primary categories: 'healthy' and 'unhealthy.' These categories represent the key labels for our classification task.

1) Healthy data split: Within the 'healthy' category, we further partitioned the dataset into 'train' and 'test' subsets. The 'healthy train' subset comprises a total of 1491 images. This subset is instrumental in training the model to recognize healthy tomato leaves effectively. It forms the basis for teaching the model the visual characteristics of undisturbed, healthy foliage. The 'healthy test' subset, on the other hand, consists of 100 distinct images. This collection serves as an independent assessment tool to gauge the model's performance. These images, being separate from the training data, help us evaluate how well the model generalizes to unseen healthy leaf samples.

2) Unhealthy data split: The 'unhealthy' category is a bit more nuanced. It encompasses tomato leaves affected by various forms of blight, encompassing both early and late stages of the disease. These stages are categorized as 'unhealthy' for our classification purpose. The 'unhealthy train' subset contains 2809 images. This vast and diverse collection offers a comprehensive training experience for the model to learn the intricacies of identifying blight in its various forms. The inclusion of both early and late stages ensures that the model grasps the entire spectrum of blight-related visual cues. The 'unhealthy test' subset, consisting of 100 images, is divided further into 29 images representing the early stages of blight and 71 images depicting the late stages. This differentiation helps us evaluate the model's proficiency in distinguishing between the progression of blight, which is a critical aspect of disease classification in the agricultural context.

3) Data validation during training: It's imperative to ensure that our model generalizes well and doesn't overfit to the training data. To address this concern, we implement a

<table>
<thead>
<tr>
<th>Limitation</th>
<th>Description</th>
</tr>
</thead>
</table>
| Limited Dataset Diversity | ○ Focus on specific diseases or crops.  
○ Lack of diversity may hinder model robustness. |
| Generalization to New Diseases | ○ Models might struggle with novel, unseen diseases. |
| Transferability across Crops | ○ Investigating the generalizability is essential.  
○ Models designed for one crop may not adapt well to others.  
○ Assessing cross-crop transferability is crucial. |
| Robustness to Environmental Factors | ○ Impact of environmental variations on model performance.  
○ Models may need to handle real-world, uncontrolled conditions. |
| Interpretability of Deep Learning Models | ○ Complex models like CNNs often lack interpretability.  
○ Understanding model decisions is important for user trust. |
| Data Challenges Annotation | ○ Manual annotation challenges affect dataset quality.  
○ Exploring improved annotation methods is necessary. |
| Real-time Application Challenges | ○ Computational efficiency and hardware constraints for real-time deployment.  
○ Exploring lightweight architectures and edge computing solutions. |

LeNet has been applied to the identification and classification of tomato illnesses while requiring the least amount of CPU processing power. Moreover, to increase classification accuracy, the automatic feature extraction technique has been used [32]. Table I provides a comprehensive summary of the limitations and gaps identified in current research on image processing applications for plant disease detection and classification.

The work in this paper supports the ResNet-18 architecture and data preparation techniques on a well-organized dataset. With an impressive 91% F1 score, the model demonstrates its ability to differentiate between good and unhealthy tomato leaves. The model's outstanding performance is demonstrated by metrics like accuracy, sensitivity, specificity, and a high AUC score on the ROC curve. This work is significant because it has real-world applications for early disease identification in agriculture. The ResNet-18 model is a potent instrument for crop management that contributes to sustainable agriculture and global food security because of its high precision and specificity.
validation procedure during training. Specifically, we reserve 15% of the data in each training batch for validation. This data is selected randomly in a way that ensures its independence from the training set. The model is periodically evaluated on this validation data to monitor its performance and make adjustments to the training process as necessary.

B. Data Preprocessing

Tomato leaf images, like many real-world images, exhibit natural variations due to factors such as lighting, background, and color. To enhance the model's ability to classify tomato leaves accurately, we employ several preprocessing steps.

One of the critical preprocessing steps is the application of a mild blurring filter. This filter is designed to reduce noise in the images, resulting in a cleaner and more consistent dataset. By doing so, we aim to minimize the impact of minor variations in image quality that may not be indicative of the actual health state of the tomato leaf. Fig. 1 illustrates the result of applying the blurring filter with a 5x5 kernel, highlighting the smoothing effect on the image.

In addition to blurring, we employ data augmentation techniques to augment the training dataset. Data augmentation is a strategy that introduces variability into the training data by applying random transformations to the images. This process helps the model generalize better, as it exposes the model to a wider range of scenarios and variations during training.

The data augmentation techniques used includes random rotations and flips. Random rotations introduce diversity by rotating images by various degrees, simulating the variability in the orientation of tomato leaves in real-world scenarios. Flips, both horizontally and vertically, add further diversity by reflecting images, mirroring the possible orientations of leaves. This augmentation is shown in Fig. 2. These preprocessing and data augmentation steps collectively serve to enhance the robustness of our classification model. They enable the model to better handle the inherent variations in real-world images, leading to improved accuracy and generalization in the task of tomato leaf classification.

![Fig. 1. Result of blurring filter with a 5x5 kernel.](image1)

![Fig. 2. Result of random rotations and flips.](image2)
IV. MODEL ARCHITECTURE

In this research, we opt for the ResNet-18 architecture as the backbone for our classification model. ResNet, short for Residual Network, is a class of neural network architectures that has demonstrated remarkable effectiveness in various deep learning tasks, particularly in image classification. What sets ResNet apart is its ability to address the vanishing gradient problem, a common challenge when training deep networks, through the use of skip connections.

ResNet-18 is a specific variant of the ResNet architecture that we have chosen for our task. It strikes a well-balanced compromise between model depth and computational efficiency. This balance is crucial, especially in applications where both high accuracy and efficient processing are essential. ResNet-18’s architecture is structured in a way that enables it to capture intricate features from images, making it an ideal choice for our task of tomato leaf classification. Fig. 3 depicts the Resnet 18 architecture.

ResNet-18 constitutes a convolutional neural network (CNN) architecture specifically tailored for image classification tasks. Noteworthy for its ability to accommodate input images of 224x224 pixels, this architecture serves as an exemplar in the ResNet lineage, embodying the principles of residual learning to facilitate the training of deep networks.

A. Convolutional Layers

The initial layer of ResNet-18 is a conventional convolutional layer, employing a 7x7 kernel and 64 filters. This layer is succeeded by batch normalization and rectified linear unit (ReLU) activation. Subsequently, a pivotal spatial reduction is accomplished through a 3x3 max-pooling layer with a stride of 2, optimizing the network’s capacity to process spatial features.

B. Residual Blocks

ResNet-18 features four residual blocks, each composed of two consecutive convolutional layers. Within these blocks, each convolutional layer is succeeded by batch normalization and ReLU activation. The hallmark of ResNet architecture, the inclusion of residual connections, enables the creation of shortcut paths, allowing information to bypass one or more layers. This strategic integration mitigates the vanishing gradient problem, empowering the network to learn more effectively.

C. Global Average Pooling (GAP)

Following the residual blocks, a Global Average Pooling (GAP) layer is applied. This layer computes the average value of each feature map, ensuring a fixed-size output independent of the input dimensions. This pooling operation contributes to the model’s spatial invariance and parameter reduction, paving the way for more efficient processing in subsequent layers.

D. Model Training

To prepare our model for the specific task of classifying tomato leaves as healthy or unhealthy, we undertake a comprehensive training process. Our model is trained for a total of 100 epochs. The choice of training duration is influenced by the complexity of the problem and the size of our dataset.

Fig. 3. Proposed model architecture with ResNet-18 classifier.
E. Objective Function Binary Cross-Entropy Loss

In this binary classification problem, we employ Binary Cross-Entropy (BCE) loss as our objective function. BCE loss is a standard choice for binary classification tasks and is well-suited for distinguishing between healthy and unhealthy tomato leaves. It quantifies the dissimilarity between the predicted class probabilities and the ground truth labels. The equation for BCE Loss is shown in Eq. (1) where \( N \) represents the total number of data points or samples in your dataset. In the context of the tomato leaf classification, each image of a tomato leaf (whether healthy or unhealthy) is considered a data point. \( y_i \) is the ground truth label for the i-th data point. In binary classification, this is a binary value indicating the actual class of the data point. For instance, \( y_i = 1 \) might represent an unhealthy leaf, and \( y_i = 0 \) might represent a healthy leaf. \( \hat{y}_i \) is the predicted output or probability assigned by the model for the i-th data point. In the context of binary classification, this value represents the model's estimate of the probability that the given data point belongs to the positive class (unhealthy) or negative class (healthy).

\[
BCE = -\frac{1}{N} \sum_{i=0}^{N} y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i) \tag{1}
\]

We utilize the AdamW optimizer during training, setting the initial learning rate at 1e-4. AdamW is a variant of the Adam optimizer that introduces weight decay, contributing to improved training stability. To further optimize training, we incorporate the Cosine Annealing learning rate scheduler with a \( T_{max} \) (maximum number of iterations) of 10. This scheduler cyclically adjusts the learning rate, allowing the model to explore different regions of the loss landscape. While a warmup learning rate scheduler was initially implemented during the early training iterations, it was eventually discarded due to providing minimal benefit.

F. Early Stopping

To prevent overfitting and to make the training process more efficient, we implement early stopping. This technique introduces a ‘patience’ parameter, set at 5 epochs in our case. The validation loss serves as the key metric for early stopping. If the validation loss does not improve for five consecutive epochs, the training process is halted.

G. Model Selection

Model selection is a critical step in our training process. We choose the top five models with the lowest validation loss and assess their performance. Ultimately, the model with the highest validation F1-score among these top five models is selected as our final model. The F1-score is particularly valuable in binary classification tasks as it balances precision and recall.

H. Model Testing

The chosen model is subjected to a comprehensive evaluation process using a range of metrics. These metrics include the F1-score, accuracy, sensitivity, and specificity. The use of multiple metrics allows us to assess the model's performance from different angles and provides a more comprehensive view of its capabilities.

I. ROC Curve and Operating Point

To determine the model's operating point, we leverage the Receiver Operating Characteristic (ROC) curve. The ROC curve illustrates the trade-off between sensitivity and specificity at various threshold settings. By analyzing this curve, we can select an operating point that best suits the specific needs of our application. This operating point defines the decision boundary for classifying tomato leaves as healthy or unhealthy. We fine-tune the pre-trained ResNet-18 model on our tomato leaf dataset, using Binary Cross-Entropy (BCE) loss as the objective function. BCE loss is a common choice for binary classification problems, such as distinguishing between healthy and unhealthy tomato leaves.

V. RESULTS

In our study, we conducted an extensive evaluation of our tomato leaf classification model, and the results demonstrate its exceptional performance in this critical task. The performance metrics we have achieved are truly remarkable and bode well for the practical application of the model. First and foremost, our model exhibited an F1 score of 91%, which is a noteworthy composite metric combining both precision and recall. This F1 score reflects the model's robust ability to accurately classify both healthy and unhealthy tomato leaves. Moreover, our model's accuracy reached an impressive 97%, indicating its proficiency in correctly categorizing tomato leaves.

Precision, which measures the ratio of true positive predictions to the total positive predictions, is a vital metric in binary classification problems. In our case, the model achieved a precision of 90.2%, highlighting its ability to make accurate positive predictions. Additionally, the model showed an outstanding sensitivity of 92%, underscoring its capability to correctly identify unhealthy tomato leaves. Notably, the specificity of the model was also high, at 90%. This indicates that the model was effective in correctly identifying healthy tomato leaves, reducing the risk of false alarms in disease detection systems.

The performance metrics depicted in Table II and Fig. 4 summarizes the evaluation of the two-class classification model for tomato leaf health. The F1 score, a balanced measure of precision and recall, stands at an impressive 91%, indicating the model's robustness. With 97% accuracy, 90.2% precision, 96% sensitivity, and 98% specificity, the model demonstrates high proficiency in distinguishing between healthy and unhealthy tomato leaves, showcasing its reliability in practical agricultural applications.

<table>
<thead>
<tr>
<th>PERFORMANCE METRIC</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 score</td>
<td>91%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>97%</td>
</tr>
<tr>
<td>Precision</td>
<td>90.2%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>96%</td>
</tr>
<tr>
<td>Specificity</td>
<td>98%</td>
</tr>
</tbody>
</table>
Fig. 4. Metrics for the two-class classification problem.

The visual representation of our model's performance in the confusion matrix (see Fig. 5) offers a more detailed breakdown of its classification accuracy, illustrating the number of true positives, true negatives, false positives, and false negatives. Furthermore, the Receiver Operating Characteristic (ROC) curve, depicted in Fig. 6, plays a crucial role in binary classification tasks. It assesses the trade-off between the true positive rate and the false positive rate at various classification thresholds. Our model's substantial Area Under the Curve (AUC) score of 0.92 on the ROC curve is a testament to its ability to effectively balance the need for disease detection while minimizing the risk of false alarms.

These results underscore the significant potential of our model for practical applications in agriculture. The high F1 score, accuracy, sensitivity, specificity, and AUC, coupled with the model's capacity to balance precision and recall, make it a powerful tool for early disease detection and crop health management. In conclusion, our research findings indicate that our classification model is highly effective and has promising implications for the agriculture industry. It provides a valuable and reliable tool for the early identification of tomato leaf diseases, which can profoundly impact crop yield, food security, and the overall health of the global food supply chain.

The figures, presented as Fig. 7 and Fig. 8, visually distinguish a healthy leaf (see Fig. 7) from an unhealthy one (Fig 8). These depictions serve to provide a quick and clear reference, enabling visual recognition of key characteristics associated with leaf health and distress.
VI. DISCUSSION

The remarkable performance of our ResNet-18-based classifier in the task of tomato leaf classification opens up significant avenues for real-world applications, particularly in the field of agriculture. The model's capacity to distinguish between healthy and unhealthy tomato leaves with over 90% accuracy, sensitivity, and specificity is not only a testament to its efficacy but also holds substantial promise for addressing real-world agricultural challenges, especially in the context of early disease detection and mitigation. One of the most noteworthy implications of our results is the potential for early disease detection. The ability to accurately identify unhealthy tomato leaves with a high degree of sensitivity allows farmers and agricultural professionals to take swift and targeted actions. Early detection of diseases such as blight can lead to more efficient intervention strategies, reducing crop losses, and minimizing the need for extensive pesticide application. This not only has financial benefits for farmers but also contributes to more sustainable farming practices by reducing the environmental impact of pesticide usage.

The high precision of our model, with a precision score of 90.2%, is a critical aspect of its practical utility. It implies that when our model makes a positive prediction (i.e., classifies a leaf as unhealthy), it is overwhelmingly likely to be accurate. This reliability is paramount when it comes to making decisions about crop management and implementing disease control measures. Farmers can have confidence in the model’s assessments and act promptly to protect their crops. Equally significant is the specificity of our model, which, at 90%, demonstrates its ability to correctly identify healthy tomato leaves. This means that the risk of false alarms, where healthy plants are incorrectly identified as unhealthy, is minimal. Again, this aspect is crucial in practical agricultural applications where false alarms can lead to unnecessary actions and resource allocation. The high Area Under the Curve (AUC) score obtained on the Receiver Operating Characteristic (ROC) curve is of paramount importance. It signifies the model's proficiency in distinguishing between healthy and unhealthy tomato leaves while maintaining a low rate of false positives. In practical terms, this means that our model effectively strikes a balance between the need for disease detection and the avoidance of unnecessary interventions. The substantial AUC score reassures farmers and agricultural professionals that the model's predictions are both accurate and reliable.

VII. CONCLUSION

In summary, our research showcases the immense potential of deep learning and convolutional neural networks in addressing pressing challenges in agriculture, with a specific focus on early disease detection in tomato crops. This technology is not confined to tomatoes alone and can be extended to various other crops, offering invaluable insights and support for sustainable agricultural practices. To harness this potential fully, future work should focus on the practical deployment of these models, integrating them with smart agriculture systems that enable timely responses to disease outbreaks, thus ensuring global food security and promoting sustainable agriculture.

Our observations are a reflection of the robustness of our classification model. The high accuracy, sensitivity, specificity, precision, and AUC score are the result of a well-trained model that has learned to recognize the subtle visual cues associated with healthy and unhealthy tomato leaves. The data preprocessing steps, including blurring and data augmentation, have contributed to the model's ability to handle real-world variations in leaf images. Additionally, the choice of the ResNet-18 architecture played a pivotal role in the model's success. ResNet architectures, known for their skip connections, are adept at training deep neural networks effectively. ResNet-18, in particular, struck a balance between model depth and computational efficiency, making it suitable for our classification task. The success of our model underscores the potential of AI and deep learning in addressing agricultural challenges, offering a valuable tool for farmers and researchers to enhance crop management, reduce disease-related losses, and contribute to more sustainable and efficient agricultural practices. The balance between accuracy, sensitivity, and specificity, along with the AUC score, emphasizes the model's real-world applicability, making it a promising asset for the agriculture industry.

REFERENCES

Proceedings of the 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence); Noida, India. 29–31 January 2020.


