Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

ICASET
2024

Research Article
Collection: Intelligent Computing of Applied Sciences and Emerging Trends

Plant Disease Detection and Classification Using Deep Learning

Muhammad Mudassar Qureshi[1], Muhammad Munwar Iqbal[1], Shabana Ramzan[2*], Saqib Majeed[3], and Muhammad Salman Bashir[4]

[1]Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan.
[2]Department of Computer Science & IT, Govt Sadiq College Women University Bahwalpur, Pakistan.
[3]University institute of Information Technology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan.
[4]Department of Computer Science & Information Technology, Virtual University of Pakistan, Lahore, Pakistan.
*Corresponding Author: Shabana Ramzan. Email: shabana@gscwu.edu.pk

Academic Editor: Salman Qadri **Published:** April 01, 2024

Abstract:. Tomato leaf disease identification using computer vision is a cutting-edge application of artificial intelligence (AI) and image processing techniques to agriculture. It plays a crucial role in modern farming by enabling early and accurate detection of diseases that affect tomato plants. This technology leverages the power of computer vision to automatically analyze images of tomato leaves and identify any signs of diseases, providing farmers with valuable insights and helping them take timely corrective measures. In this paper, we propose a completely automated end-to-end tomato leaf disease identification frameworks using Convolutional Neural Networks (CNN) and Computer Vision (CV). The proposed framework is trained and evaluated on Plant Village database containing several tomato leaf diseases along with healthy samples. We employed transfer learned CNN as well as developed a custom CNN architecture with extensive hyper-parameter optimization. The proposed framework obtained the highest classification accuracy of 92% on unseen samples. Hence, the system can be efficiently utilized for the task of real-time tomato leaf disease identification and classification.

Keywords: Tomato Leaf Disease; Plant Disease Detection; Deep Learning; Artificial Intelligence; Parameter Tuning; Transfer Learning; Convolutional Neural Network.

1. Introduction

Plant diseases encompass a wide range of infections that afflict various crops and plants, ultimately impacting agricultural productivity and ecosystem health. Fungal diseases, for instance, can cause devastating losses in crops like wheat, rice, and maize. Take the example of the Irish Potato Famine in the 19th century, where a potato late blight caused by the pathogen Phytophthora infestation led to a catastrophic failure of the potato crop, resulting in widespread famine and loss of life.

There are some bacterial diseases such as fire blight and citrus canker that has bad effect on the fruit trees and plants resulting in less fruit yield. The viral diseases like mosaic diseases also causes a lot of issues like reduced plant growth, changes the appearance of leaves and fruits in a bad way. This will make them less attractive for buyers as they become less edible. The impact that plant diseases causes are not just only to economic losses it also results in loss of biodiversity because of depending too much on plant resistance to disease. It can harm the ecosystem and reduced the variety of genes required to ensure food security for long-term. When focused on growing a single crop on a larger scale helps the disease to spread easily and more rapidly like what happened to banana and Panama disease. Globalization and climate change are also big factors in the spread of plant diseases. International Traded also make it easy for the disease to spread across different countries. Also changes in the climate can affect the distribution and how bad the disease gets. There are some disease that spreads a lot on warmer and wetter weather [1,2]

Tomato leaf diseases have a significant effect on both the plant heath and also on the economy. They also cause the domino effect that badly impacts the agriculture communities and markets. The tomato leaf diseases like Early Blight, Late Blight, Septoria Leaf Spot, and others harms in a way that goes beyond just the field, therefore impacting the economy. The instant side effects of tomato leaf diseases are that it will impact badly on the plant heath and reduced plant yields. As the disease gets worse there are some symptoms which stars appearing such as yellowing, curling and spotting of leaves and sometimes it even makes the leaves to fall off. When the leaves are effected the plants photosynthesis process from which the plant get essential nutrients and energy is not carried out properly. That results in smaller sizes fruits and bad quality. Therefore farmers faces big losses as the amount of crops they harvested are less than they expected resulting in less profits and the less market supply which means that the prices will be getting higher [3,4]

Moreover, the effect on fruit quality make the economic situation worse. Tomatoes harvested from the affected leaves can have a bad shape, taste and fewer nutrients. People will not be willing to buy tomatoes that doesn't look nice or tastes bad resulting in less demand for those. This drop in demand can result in a lot of unsold tomatoes that will cause financial losses not only to farmers but also for distributers, stores and to the companies which process the tomatoes later in the supply chain. This economic impact not only felt by individual farms it will affect the whole agriculture industry. When the farmers have less crops and that are not of good quality they will not be able to generate good profits therefore they will not be able to purchase new technology, grow their farms or even buy the basic thing they needed. In the rural areas that depends on farming sees less yield coming in results in fewer jobs that are related to agriculture industries [5,6]

In order to control tomato leaf diseases, famers start using more fungicides, pesticides and some other ways to get control on the diseases. But even these are not enough to save the crops from disease that will results in more financial problems. Buying of chemicals and applying the methods in order to control the disease will cost more for farming. This means it will the of growing a unit of crop will also be increased as a result it will be more difficult for farmers to generate profits. The local and global impact of the tomato leaf disease on the economy is significant. There are some regions which rely greatly on the production of tomatoes any disease outbreaks can causes a lot of disturbance to the entire supply chains. This will not only hurt's the farmers but the consumers also. If there will not be having enough production that means the prices will becomes higher because in order to meet the demand of the community they have to purchase tomatoes from other countries [2, 7].

To Handle plant diseases, we were trying different approaches. Using deep learning in order to detect tomato leaf diseases is a new and better way that will change how people were managing the diseases in agriculture. Deep learning consists of training of artificial neural networks to automatically recognize complex patterns from a large set of data making it a powerful tool to detect plant leaf diseases. Using deep learning for finding the tomato leaf disease has a lot of benefits. These models are highly efficient and accurate in detecting different types of tomato leaf diseases. They are really good at picking up difference in the leaves which the traditional methods can miss. Also, when these models are trained they can analyze the images automatically which speed up the process of detecting the disease we will not be requiring many people to manual inspections. That's really amazing the deep learning can change and learn from new data as well so that these models can stay updated to the newer diseases also. It's very helpful in farming because the diseases changes quickly so this flexibility of deep learning models is super helpful. Another benefit of using these models is that while detecting it's not going to hurt the plants these models can be used on difference devices including mobile phones and computer systems, so it can be used by all kinds of farmers regardless of the farm size. This ensures that the small-scale farmers and the big agricultural companies can get the benefits [8,9]

To find tomato leaf diseases using deep learning involves some important steps. First, we have to create a big dataset which consists of images of both healthy and diseases tomato leaves. This will help the model to identify and differentiate between different diseases, allowing them to take action promptly. Reducing the unnecessary use of chemicals and potentially bolstering crop yields. In this manner, the intersection of deep learning and agriculture not only advances productivity but also contributes to sustainable resource management and environmental preservation [8].

2. Proposed Methodology

In this paper, we present a tomato leaf disease classification method trained and validated using on publicly available dataset containing Tomato leaf images. The proposed method employs a transferred learned architecture i.e. MobileNetV2 along with a custom CNN architecture for the task of tomato leaf disease classification. This section discusses the proposed methodology in detail. The pipeline of proposed framework is depicted in Figure 1.

2.1 Database Acquisition and Preprocessing

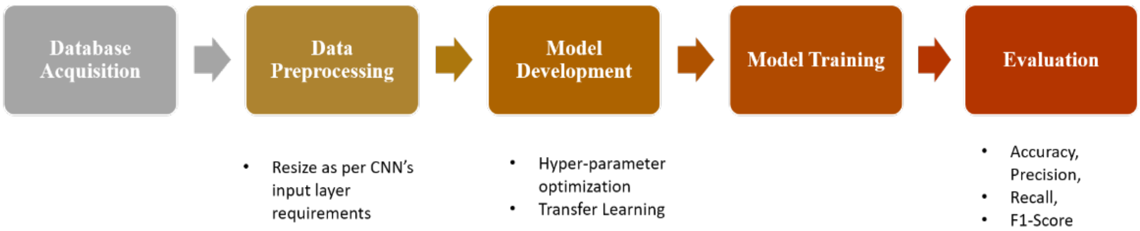

Figure 1. Proposed Framework Pipeline

The database used to conduct this study is available online. The dataset is provided by Plant Village that contains several diseased tomato leaf images along with healthy images. The images in the database are of different sizes, hence to bring uniformity, we resized the images according to input layer requirements of CNN architectures. We used 70:30 ration for training and testing. Moreover, to increase the sample sizes in database, we performed different augmentation techniques i.e. zoom, rescale and horizontal flip [10,11].

2.2 Deep Feature Extraction via CNN

After preprocessing, the features from leaf images are extracted using famous Convolutional Neural Networks. Convolutional Neural Networks (CNNs) are a type of deep learning architecture that are designed to be able to understand the visual information and process it also like the images and videos. These are inspired by human brains that how it processes the visual information. CNN comprises of layers which helps in identifying patterns and features in the data by themselves. The main part of the CNN is convolution it is a mathematical operation that uses a small filter and slides over an image to find features like edges, corners and textures. With the time these filters become good at identifying patterns. As the network goes deeper it will start combining the patterns in order to understand more complex features, helping in recognizing the objects better. CNN brought changes to various fields including computer vision, by letting the machines performs a task that was difficult to traditional algorithms. They are used in image classification to identify which objects are present in and image and object detection for finding and pointing out the different objects in an image also image segmentation where each pixel is labeled with the object it belongs to. CNNs are very efficient at finding features from the images because they are created specifically for the visual data. They are created in a way so that they can automatically find and consider important patterns and features from the images. The features can be simple like edges or textures and also can be complex like shapes and objects. The images passed through different layers each looks for a different feature. These layers use a small filter that slides over the image to identify specific visual patterns [12,13].

In CNN, every layer learns to understand the data more accurately this helps the network breaking down the visual information in a structured way starting with simple to complex details. This structured process helps the CNN model to differentiate different details helping it to identify different objects. CNN therefore performs better for various tasks of feature extraction like image classification, object detection and image segmentation. Once a CNN learned from a large dataset the inner layers can be used to extract features by passing new images and keeping the record of the outputs of specific layers. The extracted features can be passed to other machine learning algorithms so that they can be analyzed and processed further. Overall, CNN has ability to automatically extract robust and effective features from raw images, thus revolutionizing image analysis and computer vision tasks, this makes them a powerful tool for feature extraction [14,15].

In this research, we are using transfer learning to get better classification. In transfer learning we take a model that already exists and trained to do a job similar to what we need. By using the knowledge, the model learned before we can modify it to get improved results and performance on the new task. This approach is especially beneficial when we're dealing with a limited data or when we want to avoid the high computational costs associated with training a model from the beginning. Using transfer learning can significantly reduce the time and resources required to achieve satisfactory results for our classification task. Hence, is widely used by researchers to perform a task of their choice. Pre-trained models have learnt generic features from different data, which can make a model better at understanding new, unseen data. We used architectures with varying depths and complexities to extract a robust range of features from the images. These features comprise of edges, color, boundaries and pixel correlation, etc. With shallow architecture abstract features are extracted whereas, with the increase in depth, the complex feature extraction is performed [16,17]. Hence, we used deep architectures to extract a robust range of feature vectors for accurate classification. In addition to transfer learning, we also developed a custom CNN architecture for the task of tomato leaf disease classification. Several experiments are conducted to find best set of parameters and layers. The proposed architecture is illustrated in Fig. 3. The proposed CNN architecture inputs a preprocessed image of size [256 256 3] which is then supplied to a Convolution layer with 16 filters of size [3 3]. Next max pooling layer processes the image by reducing dimensionality. Again, a convolution and max pooling is applied to extract features are reduce dimensionality of image. The features are collected via dense layer and a dropout of 0.2 is applied. Final classification is performed via Softmax activation function.

3. Results and Discussion

In this paper, we propose a deep learning-based tomato leaf disease classification method using deep learning techniques. First, the images of the leaves are preprocessed, then they're analyzed by deep learning methods to find patterns and classify the diseases. This section discusses the results obtained from proposed method in detail.

3.1 Evaluation Parameters

The performance of proposed method is validated using Accuracy, Precision, Recall and F1-Score. We validated the performance of proposed framework via metrics other than accuracy due to class imbalance issues. F1-score is a powerful metric that combines the features of both precision and recall in a single metric and is widely useful in case of class imbalance issues. The metrics can be calculated as follows:

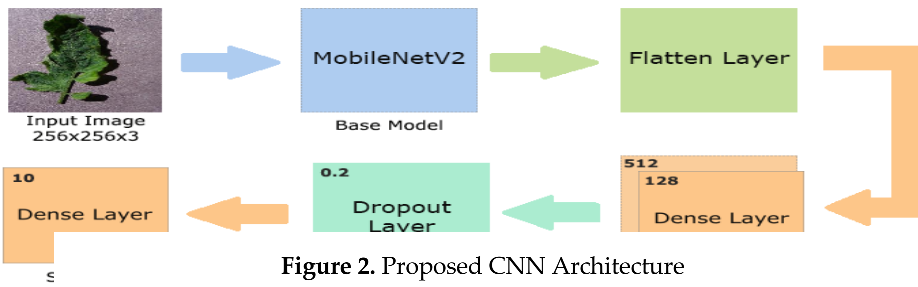

Figure 2. Proposed CNN Architecture

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

Where TP, TN, FP, FN denote True Positives, True Negatives, False Positives and False Negatives respectively.

3.2 Proposed Method Results

The proposed method obtained highest accuracy of 92% via MobileNetV2 [18]. Moreover, the custom CNN architecture obtained an accuracy of 84.2%. The results from proposed system are discussed in detail.

3.2.1 MobileNetV2

The MobileNetV2 model is utilized as the base model, with its weights set to those pretrained on ImageNet. The input shape is defined as (256, 256, 3) representing an image with a resolution of 256x256 pixels and three color channels (RGB). Following the base model setup, a custom classification head is added on top. This head consists of a Flatten layer, which converts the 3D output from the MobileNetV2 base model into a 1D array. After that, we add two layers with a type of activation function called ReLU. The first layer has 512 units, and the second one has 128 units. We also include a layer called Dropout, which helps prevent overfitting by randomly removing some connections between the layers at a rate of 0.2. Subsequently, we add a Dense layer, which uses a softmax activation function. This layer produces the final output, which consists of 10 units, each representing one of the classes in the classification task.

The suggested approach was successful in reaching the accuracy of 92%. The learning curves obtained from the architecture are illustrated in 3. The learning curves show the performance of architecture in terms of accuracy and loss. These metrics aid in determining the actual performance of architecture over each epoch. The curves indicate that the performance of proposed framework increase after 1st epoch. The training accuracy improved drastically as the model stated learning from data, whereas validation accuracy varied at different epochs and touched 90% multiple times during training. Similarly, the loss decreased after 1st epoch and the system successfully attained a 92% overall accuracy, thus making correct predications on unseen data.

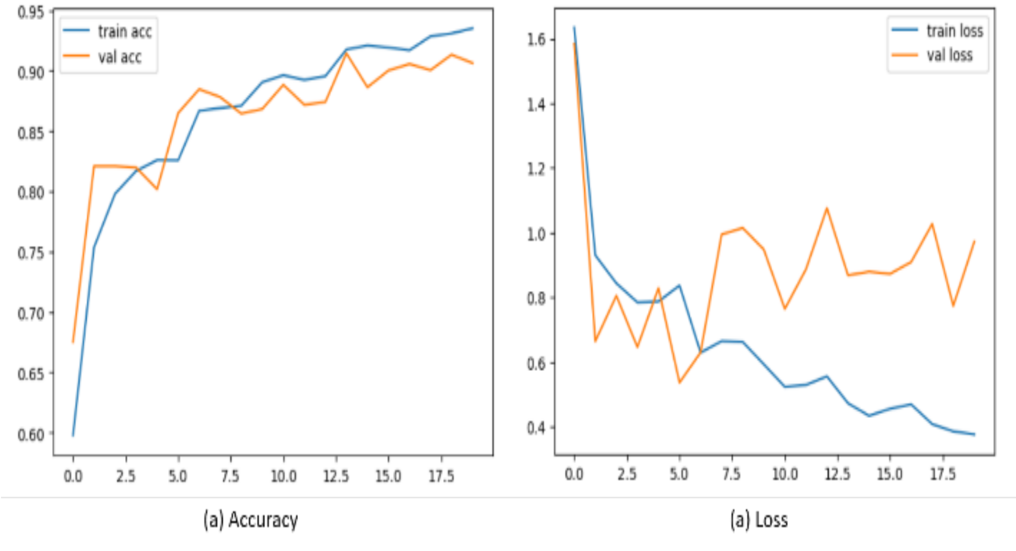

Figure 3. Learning Curves Obtained from MobileNetV2

Performance of proposed model in terms of precision, recall and f1-score is depicted in Table I. To mention, accuracy alone as an evaluation metric is not helpful in determining complete performance of classification system. Hence, determining results in terms of precision, recall and f1-score are very crucial especially in class imbalance problems. We evaluated these metrics to showcase performance of proposed framework for the task of tomato leaf disease classification. The results show that suggested MobileNetV2 framework attained precision, recall and f1-scores of 0.93, 0.88 and 0.92 respectively. The results are a proof of proposed system's efficacy and robustness. Hence, the system can be utilized in real world scenario for tomato leaf disease identification to yield crop growth.

Table I. Classification Report Obtained from MobileNetV2

Metric	Value

Precision	0.93
Recall	0.88
F1Score	0.90

The Figure 4 shows prediction of MobileNetV2 on unseen samples. The system successfully classified different diseases from images. Each image shows actual class, identified class and confidence score of prediction. The system successfully classified random samples accurately.

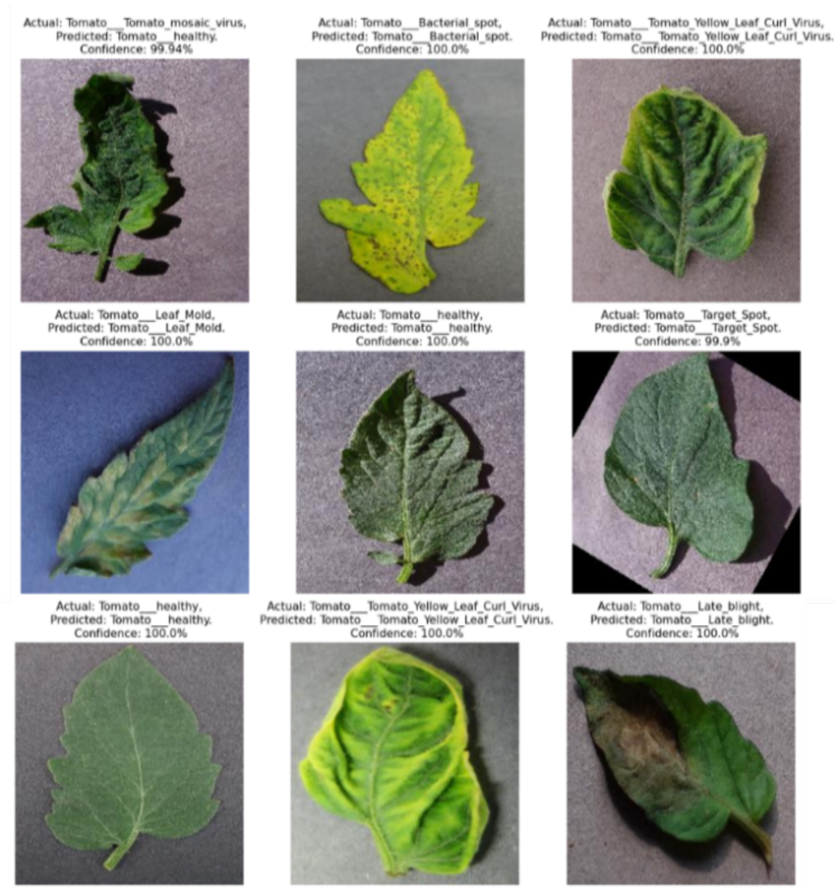

Figure 4. Prediction on Unseen Data

3.2.2 Custom CNN

We also propose a novel CNN architecture to classify tomato leaf diseases from images. This section discusses the results obtained from novel CNN architecture. The suggested framework obtained validation accuracy of 84.4% for the task of tomato leaf disease identification from leaf images. The performance in terms of accuracy and loss on train test sets is depicted in Table 2.

Table 2. Custom CNN Results In Terms of Train and Test

Split	Accuracy	Loss
Train	91%	0.26
Test	84.4%	0.48

The learning curves obtained from custom CNN architecture are depicted 5. The curves show that the proposed framework performed much better while making predictions. We reached the concluded architecture and hyper-parameters after extensive experimentations. Hence, the proposed framework is very robust and effective for tomato leaf disease identification and classification.

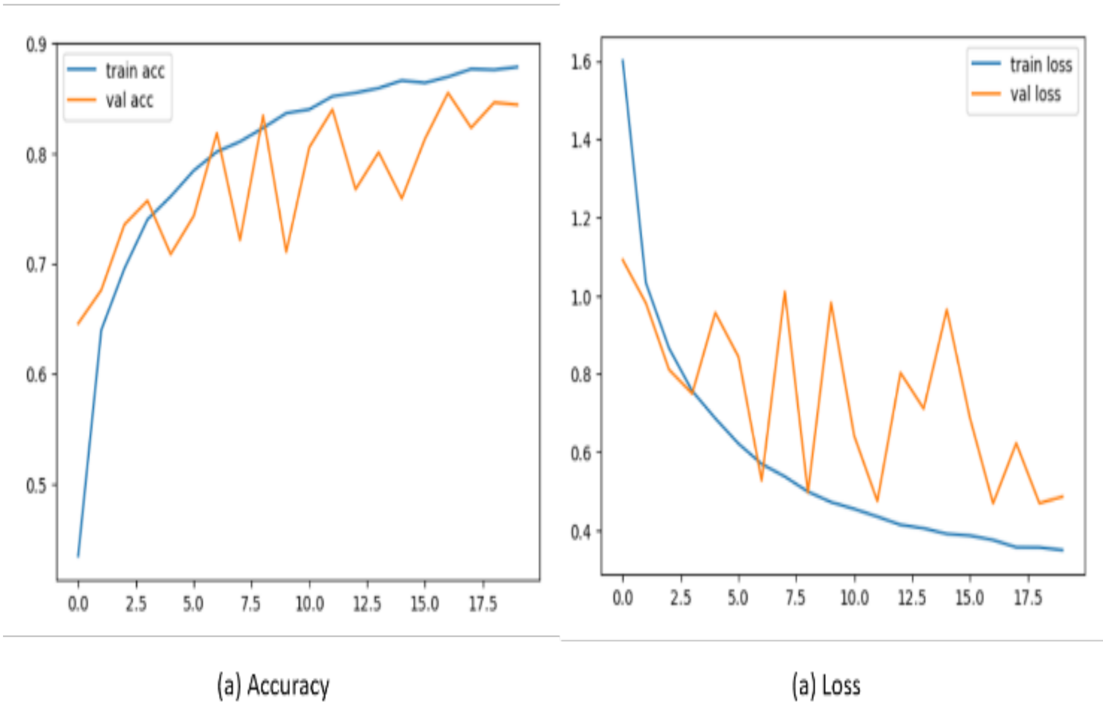

Figure 5. Learning Curves Obtained from Custom CNN.

The prediction performance of proposed framework on unseen validation samples is depicted in in Figure6. The novel CNN architecture classified most of the samples from unseen validation set accurately, thus showing performance of presented framework despite issues like class imbalance, poor contrast/illumination and background pixels in validation set. The results in terms of precision, recall and f1-score are computed in

Table3. The presented CNN architecture obtained precision, recall and f1-score of 0.81 respectively. The results depict that the system is effective for real-time tomato leaf disease identification and classification, hence, can be used for real-time detection.

Table 3. Classification Report from Custom CNN

Metric	Value
Precision	0.81
Recall	0.81
F1-Score	0.81

3.3 Parameter Tuning

Parameter optimization is an important part of deep learning in which multiple model parameters are fine-tuned to achieve the best potential outcomes. The advantages of parameter optimization in deep learning are numerous. For starters, it has the potential to increase model performance. Experimenting with various parameters will help you find the optimal balance that generates the most accurate predictions. Hence tuning these parameters is critical in order to build the best model feasible.

Now we will discuss about the parameter tuning we have done to our custom CNN architecture. We run a lot of tests in order to find the best parameters on which the model is performing better for tomato leaf disease identification. These parameters are discussed in

Table 4, whereas final parameters are highlighted in bold text.

Table 4. Parameter Optimization

Parameters	Values
Epochs	10, 20, 30
Learning Rate	0.01, 0.001, 0.000
Batch Size	8, 16, 32
Optimizer	SGD, RMSprop, Adam

Image Size	64, 128, 256

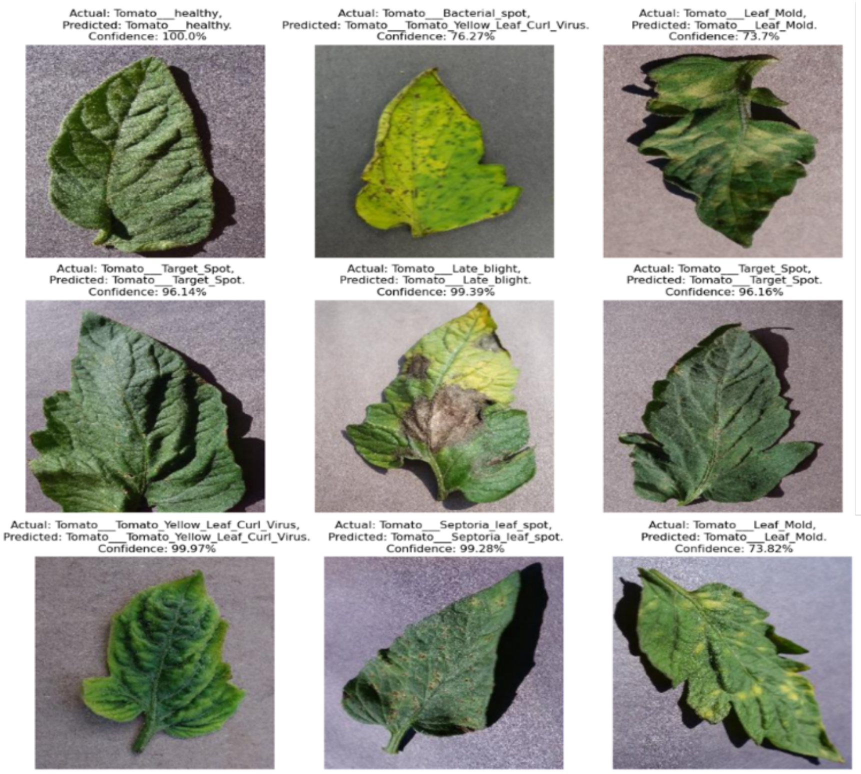

Figure 6. Prediction of Custom CNN on Unseen Samples

4. Conclusion and Future Work

Deep learning has a lot of practical applications in agriculture and plants health management, including plant leaf disease detection. This technology helps farmers and agricultural experts to quickly detect the disease that was affecting the plant leaves and respond to it accordingly by using machine learning algorithms. It has the ability to help in improving the agricultural productivity and reduces the financial losses caused by the crop diseases. Additionally, it reduces the need of pesticides because the treatment is done on the targeted area only which is also better for the environment and increases food security making it essential for agricultural methods. Other than agriculture this technology can be used in other fields as well like forestry, horticulture, and urban green spaces, all of these depend on early disease detection so that the plants will be healthy and vibrant. Also the monitoring is made easier as it detects the disease more quickly and thoroughly which results in strong and efficient farming. In this paper, we introduce a fully automated system for identifying and classifying tomato leaf diseases using leaf images. We have tested our system on a top-notch public database with 10 types of tomato leaf diseases. These leaf images are analyzed by CNN models to extract detailed features and classify the diseases accurately. In the proposed framework, we used MobileNetV2 as well as a custom CNN architecture with extensive hyper-parameter optimization for the task of tomato leaf disease identification and classification. The results show that our proposed model is efficient and robust, hence it can be easily used by farmers for real-time diagnosis of tomato leaf diseases. In the future, we aim to expand the database by using multiple leaf images to cover more diseases and make the system more robust and effective.

4.1 Comparison with Existing Systems

In this section, we're going to see how well our new system performs for detecting and classifying plant diseases compared to other existing systems. The presented framework employs a transfer learned architecture as well as a custom CNN architecture. Detailed comparison is provided in Table. Sujhata et al. [19] used Inception V3 that obtained 89% accuracy for the classification of plant leaf disease. Similarly, a transfer learning approach was employed by Hong et al. [11] using Shuffle Net that resulted in 83% accuracy on tomato leaf disease detection and classification. A CNN based approach was proposed by Salonki et al. [20] that obtained a 91% accuracy. However, they detect only tomato Spotted Wilt disease. On the contrary, the presented framework obtained a 92% accuracy that detects and classifies a total of 10 different tomato leaf diseases. In this paper, we propose a custom CNN architecture as well as use transfer learning architecture to detect and classify diseases from tomato leaf images. The presented framework is robust and effective and can be used by farmers and researchers for real-time detection of diseases in tomato leaf images.

Table 5. Comparison with Existing Systems

Reference	Technique	Performance
Sujatha et al., 2021	Inception V3	89%
Hong, Lin, & Huang, 2020	Shuffle Net	83%
Salonki et al., 2021	CNN	91 %
Proposed	**MobileNetV2**	**92%**

References

1. Agarwal, M., Singh, A., Arjaria, S., Sinha, A., & Gupta, S. J. P. C. S. (2020). ToLeD: Tomato leaf disease detection using convolution neural network. 167, 293-301.
2. Rahman, S. U., Alam, F., Ahmad, N., Arshad, S. J. M. T., & Applications. (2023). Image processing based system for the detection, identification and treatment of tomato leaf diseases. 82(6), 9431-9445.
3. Ahmad, I., Hamid, M., Yousaf, S., Shah, S. T., & Ahmad, M. O. J. C. (2020). Optimizing pretrained convolutional neural networks for tomato leaf disease detection. 2020, 1-6.
4. Saeed, A., Abdel-Aziz, A., Mossad, A., Abdelhamid, M. A., Alkhaled, A. Y., & Mayhoub, M. J. A. (2023). Smart Detection of Tomato Leaf Diseases Using Transfer Learning-Based Convolutional Neural Networks. 13(1), 139.
5. Kibriya, H., Rafique, R., Ahmad, W., & Adnan, S. (2021). Tomato leaf disease detection using convolution neural network. Paper presented at the 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST).
6. Attallah, O. J. H. (2023). Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection. 9(2), 149.
7. Agnihotri, S., Gupta, J., Garg, N., & Khatri, P. (2023). Comparative Analysis of Tomato Leaf Disease Detection Using Machine Learning. Paper presented at the 2023 6th International Conference on Information Systems and Computer Networks (ISCON).
8. Guerrero-Ibañez, A., & Reyes-Muñoz, A. J. E. (2023). Monitoring Tomato Leaf Disease through Convolutional Neural Networks. 12(1), 229.
9. Seth, V., Paulus, R., Kumar, A. J. I. S., & ICISSI, S. I. P. o. (2023). Tomato Leaf Diseases Detection Using Deep Learning—A Review. 118.
10. Liu, J., & Wang, X. J. F. i. p. s. (2020). Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network. 11, 898.
11. Hong, H., Lin, J., & Huang, F. (2020). Tomato disease detection and classification by deep learning. Paper presented at the 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE).
12. Dara, S., & Tumma, P. (2018). Feature extraction by using deep learning: A survey. Paper presented at the 2018 Second international conference on electronics, communication and aerospace technology (ICECA).
13. Chohan, M., Khan, A., Chohan, R., Katpar, S. H., Mahar, M. S. J. I. J. o. R. T., & Engineering. (2020). Plant disease detection using deep learning. 9(1), 909-914.
14. Islam, M. R., & Nahiduzzaman, M. J. E. S. w. A. (2022). Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble based machine learning approach. 195, 116554.
15. Xin, M., Wang, Y. J. E. J. o. I., & Processing, V. (2019). Research on image classification model based on deep convolution neural network. 2019, 1-11.
16. Kibriya, H., Abdullah, I., & Nasrullah, A. (2021). Plant disease identification and classification using convolutional neural network and SVM. Paper presented at the 2021 International Conference on Frontiers of Information Technology (FIT).
17. Weiss, K., Khoshgoftaar, T. M., & Wang, D. J. J. o. B. d. (2016). A survey of transfer learning. 3(1), 1-40.
18. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. J. a. p. a. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications.
19. Sujatha, R., Chatterjee, J. M., Jhanjhi, N., Brohi, S. N. J. M., & Microsystems. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. 80, 103615.
20. Salonki, V., Baliyan, A., Kukreja, V., & Siddiqui, K. M. (2021). Tomato spotted wilt disease severity levels detection: a deep learning methodology. Paper presented at the 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN).