International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 03 | Mar 2024 www.irjet.net p-ISSN: 2395-0072

Identification of Tomato Leaf Disease Using Convolutional Neural Networks - UNet

[1]Ankush Nandi, [2]G.V. Anil, [3]K. Gnana Likhitha, [4]Ashmi Raj

[1,2,3,4]Student, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.

Abstract - Plants play a crucial part in preserving the ecosystem's equilibrium and supplying the resources required for human survival. Tomatoes, scientifically known as Solanum lycopersicum, are among the most widely cultivated and consumed vegetables globally. They are a good source of potassium, folate, and vitamin C, among other important minerals and vitamins. This nutritional value makes them a valuable cropfor ensuring food security and promoting a healthy diet. Despite their significance, tomato crops face several challenges, with diseases being one of the major issues. The emergence of diseases can lead to substantial yield losses and quality deterioration, posing a significant threat to both local and international tomato markets. Timely and accurate disease detection is crucial to mitigate these losses and maintain the economic viability of tomato cultivation. Our research paper addresses the crucial problem of tomato leaf disease detection in this particular context. To automatically identify tomato leaf diseases like bacterial spot, late blight, early blight, spider mites, mosaic virus, yellow leaf virus, target spot, and septoria leaf spot, we suggested a novel method. utilizing the U-Net architecture and convolutional neural networks (CNN). Our study makes use of a dataset that we gathered from multiple sources that includes pictures of both healthy and unhealthy tomato leaves. Based on the research, when compared with VGG16, VGG19, CNN achieved higher accuracy. So, we are trying to employ a CNN for feature extraction and a U-Net for pixel-wise segmentation, allowing us to pinpoint disease-affected areas with high precision. We anticipate that the suggested method's processing time will be much less than that of these models.

1. Introduction

For thousands of years, agriculture has been the main source of human survival. Every year, plant diseases cause significant losses in crop productivity across the world. A vital role in the economy is played by agriculture, the threat of crop diseases casts a long shadow over both crop quality and quantity, but they are prone to a number of illnesses that could seriously harm them and reduce their productivity. The promptidentification and precise diagnosis of diseases affecting plant leaves is one of the main obstacles in the management of plant diseases.

The field of agriculture could be strengthened by automating the detection of leaf disease using image processing techniques like deep learning and machine learning, which could yield quickand accurate results. These methods entail a number of processes, such as feature extraction, machine learning, segmentation, processing, and image acquisition. Plant leaf diseases can be identified and categorized from visuals using machine learning techniques. Accurate disease detection and classification are made possible by the utilization of machine learning algorithms to extract features from which patterns and relationships can be learned

Machine learning is used to identify and classify various fungal, bacterial, and viral diseases in plant leaves. These diseases can significantly affect crop yield and quality. Different machine learning algorithms are employed for disease classification. These comprise artificial neural networks (an ANN), support vector machines, and deep learning models such as GoogLeNet and the network CNN. SVM is used to classify diseases into multiple classes using the features that are extracted. RGB images of leaves are used as data sources for machine learning. These images contain visual information about the symptoms and characteristics of the diseases. Deep learning models are trained to automatically learn relevant features from the data, which can be particularly effective for image-based tasks.

Deep learning techniques, such as convolutional neural networks, have shown excellent performances in various image analysis tasks, including leaf segmentation, leaf spot detection, and disease classification. CNNs are designed to automatically learn hierarchical features from images. They consist of multiple layers, each responsible for detecting increasingly complex patterns. This hierarchical feature learning allows CNNs to capture relevant information at different scales, making them highly effective in recognizing intricate details in images. Because CNNs are translation-invariant, they can recognize features no matter where they are in the picture. This property is especially valuable when dealing with images of plant leaves, where the position and orientation of disease symptoms may vary. Ferentinos developed CNN models to identify and diagnose plant diseases using leaf photos of normal and infected plants.

The approach involved training the models on a dataset of leaf images and using them to classify new images as either normal or infected. They utilized the AlexNet and VGG16 models with optimized hyper parameters for disease detection in tomato crops. The study demonstrated the superiority of deep learning models over traditional machine learning techniques in accurately identifying and classifying crop diseases. It draws attention to the drawbacks of conventional manual inspection techniques, such as their subjectivity, labor intensity, and inability to scale.

2.Literature Survey

[1] Seyed Mohamad Javidan and Ahmad Banakar conducted a study that focused on utilizing K-means clustering and machine learning techniques for the diagnosis of diseases in grape leaves.

This research incorporated the use of Support Vector Machines (SVMs) and a novel image processing method to identify grape leaf ailments, including leaf blight, black rot, and black measles. By applying K-means clustering, the study extracted features from both RGB and HSV color models to differentiate between healthy and diseased regions.

To reduce dimensionality, Principal Component Analysis (PCA) was employed, followed by SVM classification, achieving a remarkable accuracy of 98.97%. Furthermore, deep learning methods, including CNN and GoogleNet, were explored, albeit consuming more computational resources and time, yielding accuracies of 86.82% and 94.05%, respectively.

[2] “Plant Leaf Disease Detection and Classification Based on Convolutional Neural Network," authored by B. Swetha and Ravindar Eslavath, delves into the application of the CNN algorithm [10] that plays a pivotal role in this research by extracting valuable features from images through the utilization of convolutional and pooling layers for the detection of plant diseases.

[3] In their paper titled "Modified U-Net for plant diseased leaf image segmentation," Shanwen Zhang and Chuanlei Zhang introduced an improved U-Net architecture aimed at enhancing the segmentation of images depicting plant leaves affected by diseases. The conventional U-Net models faced challenges when dealing with the efficient segmentation of complex crop diseased leaf images so they introduced the Modified U-Net (MU-Net) model, which incorporates a residual block and a residual path into its architecture.

[4] In the paper titled "Machine Learning-Based Tomato Leaf Disease Diagnosis Using Radiomics Features," authored by Faisal Ahmed, Manju Rahi, Raihan K. Uddin at. al, analyzed a machine learning approach for the identification and classification of tomato leaf diseases [9]. In addition to distinguishing healthy leaves, the authors have categorized diseases into three groups: Septoria, Yellow Curl Leaf, and Late Blight. The research methodology involves the extraction of features from tomato leaf images using radiomics-based features and then utilized with a Gradient Boosting Classifier enabling accurate identification and categorization of tomato leaf disease.

3. Project Goal

The goal of this research is to

1. Provide effective techniques to precisely identify tomato plant diseases at an early stage, potentially leveraging artificial intelligence tools such as machine learning and image processing.
2. To enhance crop management strategies, use contemporary technologies like CNN in image processing for disease detection.
3. By facilitating prompt detection and intervention techniques, strive to lessen the financial impact of plant diseases.

4.Problem Identification

The prevalence of diseases that affect tomato plants presents serious challenges for the agricultural sector in India. The conventional technique of visually diagnosing plant diseases in rural India is ineffective and frequently results in significant yield losses. Furthermore, the nation lacks the necessary infrastructure to effectively combat these diseases. In a similar vein, leaf diseases seriously jeopardize crop productivity and economic stability in rural India, even though the region is a major producer of tomatoes

5. Dataset Description

The information for our work was obtained from the online resource Kaggle. The dataset for Plant Village was gathered from (Gomaa, 2023) and was utilized for this project.

Over 54,000 high-quality photos of 14 different crop species, including tomato, potato, apple, and grape, have been included in the Plant Village collection.

Each image is connected to one of 38 classes, which stand for different plant diseases or healthy conditions. Our main target is on tomato leaf and the dataset of tomato has the following as shown in Table 1.

S.No	Name	Number of Images
1	TOMATO_healthy	1,273
2	TOMATO_mosiac_virus	229
3	TOMATO yellow leaf virus	4,286
4	TOMATO_target_spot	1,123
5	TOMATO_spider_mites	1,341
6	TOMATO_ Septorialeaf_spot	1,417
7	TOMATO_ Leaf Mold	761
8	TOMATO_ Late blight	1,527
9	TOMATO_ Early blight	800

Table 1: Dataset

6.Proposed Architecture

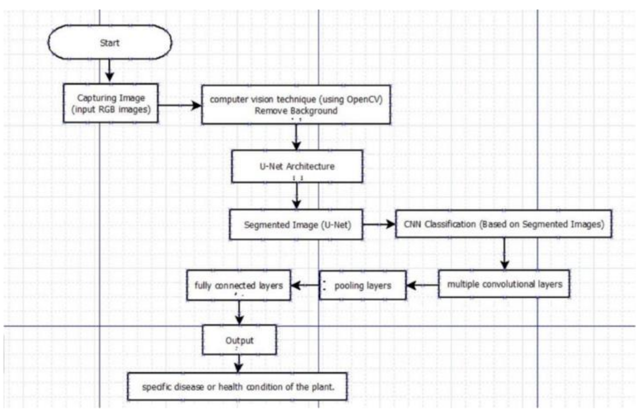


Fig 1: Our Proposed Architecture

6.1 Technical Indications

Our proposed architecture in the image Fig 1 is a computer vision system for plant disease detection. It has the following steps:

Capturing a Picture: The system is capable of taking a picture input from a number of sources, including files, scanners, and cameras. The picture must have three channels—red, green, and blue—and be in RGB format.

Remove Background: The background of the image is eliminated by the system using a computer vision technique. This helps the U-Net architecture segment the leaf more easily and helps the system concentrate on the leaf itself.

U-Net Architecture: One kind of neural network that works well for segmentation tasks is the U-Net architecture. It is made up of a decoder and an encoder. Features are extracted from the input image by the encoder. After that, the decoder uses the features that were extracted to reconstruct the image, but this time, it labels every pixel with a class label. The various leaf regions in this instance—healthy, diseased, and background—are the class labels.

CNN Classification: One kind of neural network that works well for image classification tasks is the CNN classifier. It is made up of several fully connected, pooling, and convolutional layers. The fully connected layers acquire the ability to classify the image into distinct classes, the pooling layers down sample the feature maps to lower the number of parameters and computational complexity, and the convolutional layers learn to extract features from the image. In this instance, the various segments of the segmented leaf are classified into distinct disease classes using the CNN classifier.

Output: The output of the system is the disease class of the leaf, or a prediction of whether the leaf is healthy or diseased. The system can also output a probability distribution over the different disease classes, which can be used to indicate the confidence of the system in its prediction.

Training Dataset: The test's most crucial element is the training dataset. A significant number of photos of leaves with and without illnesses ought to be included. The pictures have to show the many kinds of leaf diseases that one could potentially come across in the field.

Feature Extraction: During this phase, the model learns which leaf characteristics are most crucial for differentiating between healthy and sick leaves. Numerous feature extraction techniques are available for application. The kind of leaf diseases under investigation and the caliber of the training picture will determine which method is optimal for a certain application.

Leaf Disease Prediction: The model makes use of the extracted data to predict leaf diseases at this stage.

Leaf Disease Classification: The model classifies the type of disease if it determines that the leaf is ill. Numerous distinct leaf diseases can be classified by the model once it has been taught. The quantity and caliber of the training dataset will determine how many diseases can becategorized.

Testing Dataset: This dataset is intended to assess how well the leaf disease classification and recognition model performs.

7.Results

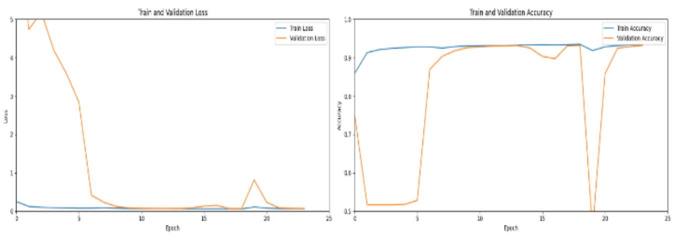
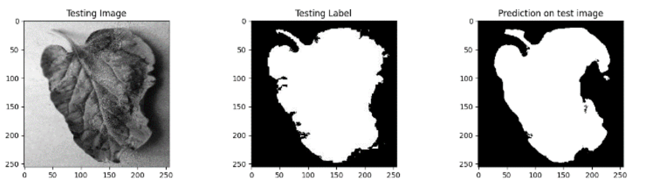


Fig 2: Accuracy



The benefit of the applied CV model in precisely identifying regions of interest within images of tomato leaves has been demonstrated by its remarkable performance metrics for the semantic segmentation of healthy tomatoes. Before the segmentation process began, plant leaf images were preprocessed by removing the background and then converted to binary segmented images that were used as mask images as shown in fig. Important training data for the U-Net model came from these mask images. The model is ready to acquire strong features necessary for accurate segmentation thanks to careful data preparation that includes image loading, resizing, and normalization. By utilizing the inherent benefits of the U-Net architecture, such as skip connections and encoder-decoder symmetry, the model is able to capture both local and global features, which can be used to accurately segment healthy tomato regions.

The model demonstrated remarkable learning capabilities during training, landing at an astounding 93% accuracy rate as shown in fig 2. This striking precision highlights the model's ability to discriminate between healthy and unhealthy tomato regions, which is important for agricultural applications that seek to monitor crops and identify diseases. At 0.9706, 0.9717, and 0.9712, respectively, the model's precision, recall, and F1-score values were also noteworthy. These measures demonstrate the model's capacity to reliably produce segmentation results by achieving a high percentage of true positive identifications while reducing false positives and false negatives. Mean Intersection over Union (IoU), a quantitative indicator of segmentation accuracy, was employed in the comprehensive assessment of the model to forecast disease regions within plant leaf images. The model's ability to correctly identify disease-affected areas was revealed by the IoU metric and a visual comparison between model predictions and ground truth masks. We were able to obtain a thorough grasp of the model's segmentation capabilities and its efficacy in assisting with disease diagnosis and management in agricultural contexts by examining the discrepancy between predicted and actual disease regions.

In summary, the success of the developed U-Net model in precisely identifying regions of healthy tomatoes is demonstrated, indicating its potential utility as a useful instrument for both researchers and agriculturalists. The model makes a substantial contribution to the fields of computer vision and agricultural technology through its careful implementation, thorough evaluation, and outstanding performance metrics.

8.Conclusion & Future Work

Our work aims to increase the efficiency in detecting tomato leaf diseases, so we suggested an approach that makes use of CNNs, which are particularly good at image analysis and classification, and U-Net, which is renowned for its skill at pixel-by-pixel segmentation, to benefit from the advantages of deep learning. We will be trying to implement this robust and accurate tool for identifying disease-affected regions in tomato leaves which might be possible by the combination of feature extraction and high-resolution segmentation. In conclusion, by utilizing image segmentation techniques and specifically the UNet architecture, our project has made significant progress toward the identification of tomato plant leaf diseases. We have achieved a remarkable 93% accuracy rate as shown in fig, exceeding the standards established by earlier research. Our approach is robust and effective, as evidenced by our precision, recall, and F1-score metrics of 0.9706, 0.9717, and 0.9712, respectively. These outcomes show how our methodology can be used to create a sophisticated and reliable tomato leaf disease detection system. To fully implement our approach, more experimental validation and optimization will be required, but the preliminary results point to a promising path towards automated and early disease detection in agriculture, laying the groundwork for more widespread applications in plant disease diagnosis.

Despite considering the notable achievements of our project, there is still much room for improvement in the field of real-time plant disease detection systems. Despite the remarkable accuracy and effectiveness of our methodology, there is a need to expedite the process for real-time implementation so that farmers and agricultural practitioners can receive immediate feedback. By enabling quick decisions based on precise and timely information, the creation of such systems has the potential to completely transform agricultural disease prevention and management. Further research efforts could concentrate on improving the suggested methodology's scalability and adaptability to suit a greater variety of plant diseases and environmental circumstances, thereby expanding its application in various agricultural contexts.

9. Reference

[1] Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning by Seyed Mohamad Javidan,AhmadBanakar,2023

[2] Zamani, Abu Sarwar, et al. "Performance of machine learning and image processing in plant leaf disease detection." *Journal of FoodQuality* 2022 (2022): 1-7.

[3] Voutos, Yorghos, et al. "A survey on intelligent agricultural information handling methodologies." *Sustainability* 11.12 (2019): 3278.

[4] Pandian, J. Arun, et al. "A five convolutional layer deep convolutional neural network for plant leaf disease detection." *Electronics* 11.8 (2022): 1266.

[5] Ferentinos, Konstantinos P. "Deep learning models for plant disease detection and diagnosis." *Computers and electronics in agriculture* 145 (2018): 311-318.

[6] Zhang, Shanwen, and Chuanlei Zhang. "Modified U-Net for plant diseased leaf image segmentation." *Computers and Electronics in Agriculture* 204 (2023): 107511.

[7] International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 11 Issue V May 2023- www.ijraset.com

[8] Talasila, S., Rawal, K. and Sethi, G., 2023. Deep Learning-Based Leaf Region Segmentation Using High-Resolution Super HAD CCD and ISOCELL GW1 Sensors. *Journal of Sensors*, *2023*.

[9] Ahmed, Faisal, Mohammad Naim Uddin Rahi, Raihan Uddin, Anik Sen, Mohammad Shahadat Hossain, and Karl Andersson. "Machine Learning-Based Tomato Leaf Disease Diagnosis Using Radiomics Features." In *Proceedings of the Fourth International Conference on Trends in Computational and Cognitive Engineering: TCCE 2022*, pp. 25-35. Singapore: Springer Nature Singapore, 2023.

[10] Sardogan, Melike, Adem Tuncer, and Yunus Ozen. "Plant leaf disease detection and classification based on CNN with LVQ algorithm." In 2018 3rd international conference on computer science and engineering (UBMK), pp. 382-385. IEEE, 2018.

[11] Raja Kumar, R., Athimoolam, J., Appathurai, A., & Rajendiran, S. (2023). Novel segmentation and classification algorithm for detection of tomato leaf disease. *Concurrency and Computation: Practice and Experience*, *35*(12), e7674.

[12] Tugrul B, Elfatimi E, Eryigit R. Convolutional neural networks in detection of plant leaf diseases: A review. Agriculture. 2022 Aug 10;12(8):1192. [13] Pandian, J. Arun, et al. "A five convolutional layer deep convolutional neural network for plant leaf disease detection." *Electronics* 11.8 (2022): 1266.

[13] Chen, Z., Wu, R., Lin, Y., Li, C., Chen, S., Yuan, Z., Chen, S. and Zou, X., 2022. Plant disease recognition model based on improved YOLOv5. *Agronomy*, *12*(2), p.365.

[14] Li, Y. and Chao, X., 2021. Semi-supervised few-shot learning approach for plant diseases recognition. *Plant Methods*, *17*, pp.1-10.

[15] Sagar S, Singh J. An experimental study of tomato viral leaf diseases detection using machine learning classification techniques. Bulletin of Electrical Engineering and Informatics. 2023 Feb 1;12(1):451-61.

[16] Jayanthi, M. G., and D. R. Shashikumar. "Eggplant leaf disease detection and segmentation using adaptively regularized multi Kernel- Based Fuzzy*C*-Means and Optimal PNN classifier."

[17] Attallah, O., 2023. Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection. *Horticulturae*, *9*(2), p.149.