iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49338

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

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 III Mar 2023- Available at www.ijraset.com

Plant-Leaf Disease Prediction Using Deep Learning

K.Ashritha[1], K.Sandhya[2], Y. Uday Kiran[3], Mr.V.N.L.N.Murthy[4]
[1, 2, 3]*Student,* [4]*Professor, Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India*

Abstract: ***Brown spot, Mosaic, Grey spot, and Rust all significantly reduce apple yield. Rust is a sign of Foliar illness in this instance. The primary factor influencing apple output is the occurrence of apple leaf diseases, which results in significant yearly economic losses. Therefore, it is very important to research apple leaf disease identification. Plants are frequently attacked by pests, bacterial diseases, and other microorganisms. Inspection of the leaves, stem, or fruit usually identifies the attack's signs. Powdery Mildew and Leaf Blight are two common plant diseases that can cause severe harm if not treated quickly. In the realm of agriculture, image processing is frequently utilized for classification, detection, grading, and quality control. Finding and identifying plant diseases is crucial, especially when trying to produce fruit of the highest caliber. The real-time identification of apple leaf diseases is addressed in this research using a deep learning strategy that is based on enhanced convolutional neural networks (CNNs). This study uses data augmentation and image annotation tools to create the foliar disease dataset, which is made up of complex images captured in the field and laboratories.***

Overall, we can identify the illness present in plants on a massive scale by utilizing machine learning to train the vast data sets that are publically available. The project explains how to identify plant leaf diseases, how they affect plant yield, and which pesticides should be used to treat them. in agriculture. To monitor huge plant fields and automatically identify disease symptoms as soon as they develop on plant leaves, research on automatic plant disease is crucial. In this essay, we'll demonstrate how to identify plant illnesses by obtaining photos of their leaves.

Keywords: Rust, CNN, Scab, Foliar Disease, Plant-Leaf

I. INTRODUCTION

The precision and dependability of detection and analysis processes are improved by the use of technology. For instance, individuals who use cutting-edge technology to study diseases that emerge suddenly have a better chance of controlling them than those who do not. India is a developed nation where agriculture supports around 70% of the people. Farmers can choose from a wide variety of eligible crops and choose the right insecticides for their plants. A considerable decrease in both the quality and quantity of agricultural products is caused by plant disease. Research on visually discernible patterns in plants is referred to as plant disease studies. In the beginning, the specialist in that sector would manually monitor and analyze plant illnesses. Techniques for image processing can be used to find plant diseases. The fruit, stem, and leaves frequently show signs of disease. The Apple plant leaf is taken into consideration for disease identification since it exhibits disease signs. This article provides an introduction to the image-processing method used to detect plant diseases. Agricultural diseases pose a serious hazard to human existence because they could trigger famines and droughts. In situations where farming is done for commercial objectives, they also result in significant losses. The diagnosis and treatment of diseases may be aided by the application of computer vision (CV) and machine learning (ML). Artificial intelligence (AI) in the form of computer vision involves utilizing computers to comprehend and recognize objects. It is presently used in medical operations to detect and evaluate items. It is principally employed in assessing drivers, parking, and driving of self-driven automobiles.

Food security is made simple by the use of computer vision to improve plant disease protection accuracy. Pests and diseases harm crops or plant parts, reducing food production and escalating food poverty. Also, little is known about diseases and pest management or control in many less-developed nations. One of the main causes of decreased food production is toxic pathogens, poor disease control, and dramatic climate change.

In recent years, disease identification methods based on servers and mobile devices have been used. Automatic disease recognition is made possible by several elements, including the high-resolution camera, high-performance processing, and numerous built-in accessories.

The accuracy of the results has been improved by using contemporary methods like machine learning and deep learning algorithms. For the detection and diagnosis of plant diseases, numerous studies have been conducted using classic machine learning techniques such as random forests, artificial neural networks, support vector machines (SVM), fuzzy logic, the K-means method, and convolutional neural networks.

©IJRASET: All Rights are Reserved

II. RELATED WORK

A. Literature Survey

1) The backpropagation neural network-based "Classification of Pomegranate Diseases" developed by S. S. Sannakki and V. S. Rajpurohit proposed primarily relies on the technique of segmenting the defective area and uses color and texture as features. For the classification in this case, a neural network classifier was applied. The key benefit is that the image's chromaticity layers are extracted using a conversion to L*a*b, and categorization is determined to be 97.30% correct. The biggest drawback is that it is only utilized for a few types of crops.
2) Hu's moments are employed as a distinguishing characteristic in "Cotton Leaf Disease Diagnosis using Pattern Recognition Methods," a method presented by P. R. Rothe and R. V. Kshirsagar. Using an active contour model to restrict the amount of energy that can enter an infection site, a BPNN classifier deals with the many class issues. The categorization rate on average is 85.52%.
3) Using computer vision technology and fuzzy logic, Aakanksha Rastogi, Ritika Arora, and Shanu Sharma published a paper titled "Leaf Disease Detection and Grading". K-means clustering is utilized to divide the defective region; GLCM is used to extract textural information; and fuzzy logic is used to grade the severity of the ailment. They employed an artificial neural network (ANN) as a classifier, which primarily aids in determining how seriously the sick leaf is affected.
4) Automatic Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease was a proposal made by Godliver Owomugisha, John A. Quinn, Ernest Mwebaze, and James Lwasa. Extracted color histograms are converted from RGB to HSV and RGB to L*a*b. The Area under the curve analysis is utilized for classification, peak components are used to form the max tree, and there are five shape attributes. Naive Bayes, Decision Tree, Random Forest, Very Random Tree, Nearest Neighbors, and SV Classifier were employed. Randomized trees produce a very high score in seven classifiers, offer real-time data, and give the application flexibility.
5) SVM-based Multiple Classifier System for Identification of Wheat Leaf Diseases by Uan Tian, Chunjiang Zhao, Shenglian Lu, and Xinyu Guo. Color features are encoded in RGB to HIS by utilizing GLCM, and seven invariant moments are used as the form parameter. They employed an SVM classifier with MCS, used for offline disease detection in wheat plants.
6) Not just in India, but many other nations also rely on agriculture as their main industry. Researchers from all over the world are using the most recent technologies to attempt and tackle various issues that farmers are currently facing. Three techniques—multi-label classification, focus loss function, and regression—have been proposed by Zhong and Zhao. They have a DenseNet121 architecture foundation. Six apple leaf diseases were present in a total of 2462 pictures of apple leaves. The proposed method outperformed more established multi-classification algorithms with an accuracy rate of 93.5%.

B. Existing System

Several procedures must be carried out to determine if the leaf is healthy or diseased. that is, Preprocessing, Feature Extraction, Classification, and Classifier Training. Preprocessing involves reducing all of the photos' sizes to a single, uniform value. The next step is to extract features from a preprocessed image with the aid of HOG.

A feature descriptor for object detection is HOG. This feature descriptor uses the gradient of its intensity to represent both the appearance of the object and the contour of the image.

The fact that HOG feature extraction uses the generated cells is one of its drawbacks. Any changes have no impact on this. Three feature descriptors were used in this case.

III. PROPOSED SYSTEM

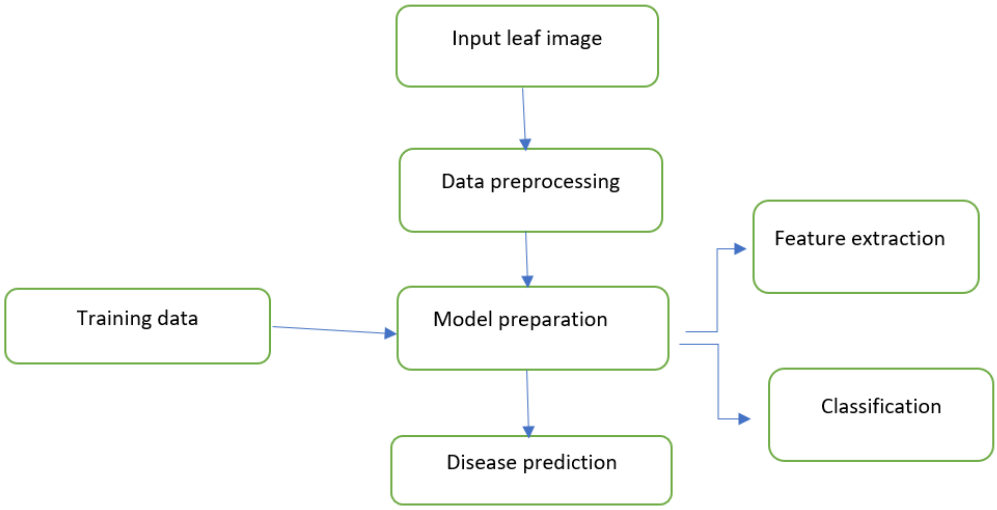
Certain procedures must be carried out to determine if the leaf is healthy or diseased. That is, Preprocessing, Feature Extraction, Classification, and Classifier Training. Preprocessing image is reducing the size of each image to a standard size. The next step is to extract features from a preprocessed image using HOG (Histogram of an Oriented Gradient). HoG is an object detection feature descriptor.

The appearance of the object and the outline of the image are characterized by the gradients in this feature descriptor. The fact that HoG feature extraction uses the newly formed cells is one of its benefits. Any changes have no impact on this.

We have all witnessed how numerous plant diseases cause plant deformation, which infects both farmers and everyone else. Due to the supply deficit, this issue has an impact on both consumer and farmers' incomes. An enormous sum of money, 60 billion dollars, is thought to have been lost as a result of plant leaf disease.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The Flow of the System:



Dataset

These are the sample images from the dataset showing symptoms of apple tree (A)Healthy leaf, (B)apple scab, (C) Rust, and (D) Multiple diseases. This dataset contains 3k pics with both train data and a test set of images.

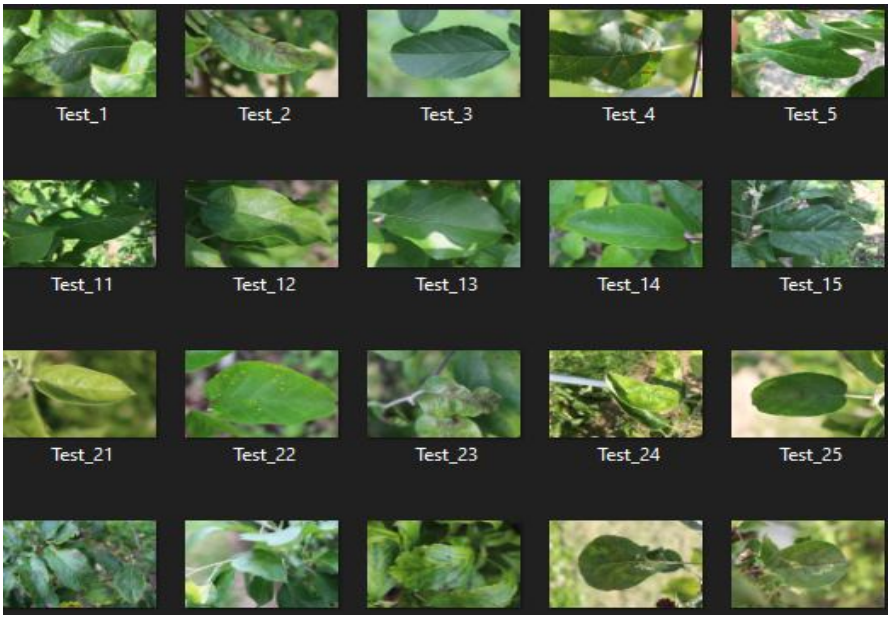


Fig 1 Dataset

The First process is to take the leaf image as input. In the next step, we do data preprocessing.

A. Data Preprocessing

Data preprocessing is a crucial step in the data analysis process, particularly in machine learning and data mining. It refers to the process of cleaning, transforming, and preparing raw data into a form that is ready for analysis. The raw data collected from various sources may contain errors, missing values, outliers, and other inconsistencies that need to be addressed before it can be analyzed. Data preprocessing involves a series of steps to prepare the data for analysis. Data preprocessing is an iterative process that may involve repeating some of these steps multiple times to refine the data and improve the accuracy of the analysis. Effective data preprocessing is critical to ensure the accuracy, efficiency, and relevance of the analysis results.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Model Preparation

The Model we have used is H5. An H5 model is a file format used to store a trained deep learning model in the Hierarchical Data Format version 5 (HDF5) file format.

When a deep learning model is trained, it generates a set of weights that capture the learned patterns and relationships between the input and output data. These weights can be stored in an H5 model file format, which can be used to later load the model and make predictions on new data. The H5 model file contains all the information necessary to recreate the trained model, including the model architecture, the values of the weights, and the training configuration used to train the model. The file can be easily shared with others, and the model can be used in different programming languages and platforms.

The HDF5 file format is a binary file format used for storing large and complex datasets, including scientific data, image data, and deep learning models. The format is designed to efficiently store and retrieve large amounts of data, making it a popular choice for storing deep learning models. To use an H5 model file, the file must first be loaded into memory using a library or framework that supports the HDF5 file format. Common deep learning frameworks such as TensorFlow, Keras, and PyTorch all have built-in support for loading and using H5 model files. Once the file is loaded, the model can be used to make predictions on new data.

C. Training Data

We train data with CNN model,

1) *CNN:* Convolutional Neural Networks are a sophisticated chain of neural networks that extract an image's attributes from a trained dataset and classify them to provide the desired result. By converting the dataset's picture data to numerical values, it trains the neural networks. The primary benefit of CNN over its predecessors is that it automatically recognizes crucial details without human supervision. ConvNets are computationally efficient and more potent than machine learning techniques. Based on their identified qualities, these numerical values are subsequently placed into numerical arrays. Depending on the input, these arrays are then placed in various network nodes and passed through many iterations. The CNN models function similarly to a filter, eliminating dust and separating the features of the images, and are utilized for geographical classification in many businesses that need data to be categorized quickly and securely.

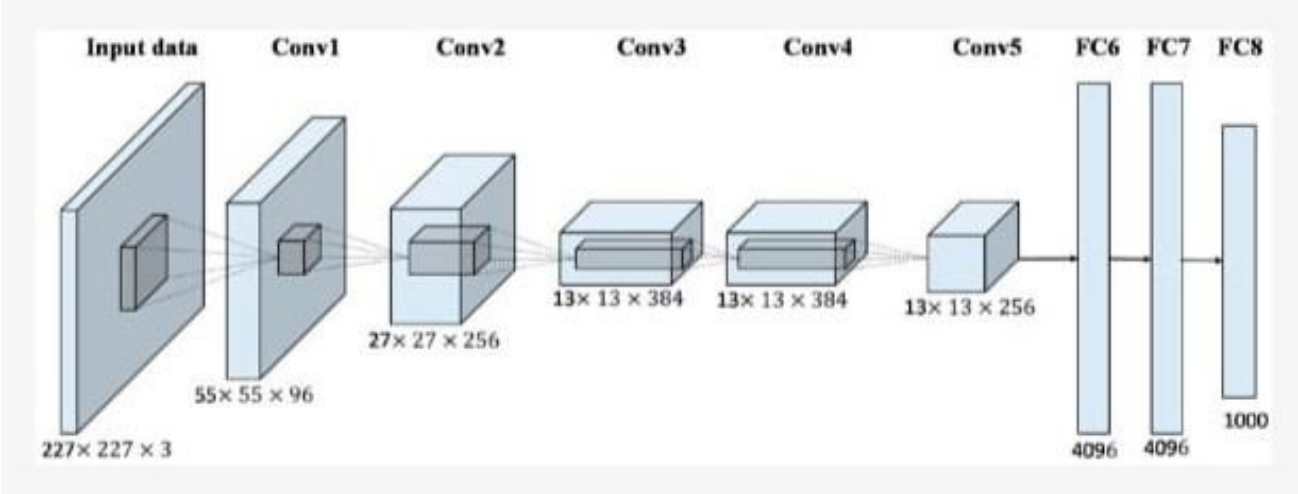


Fig 2.CNN Architecture

2) *LeNet:* LeNet, short for LeNet-5, is a convolutional neural network (CNN) architecture developed by Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Haffner in 1998. It was one of the first successful deep learning models for image recognition, specifically for handwritten digit recognition tasks. LeNet-5 consists of seven layers, including two convolutional layers, two subsampling layers, and three fully connected layers. The input to the network is a grayscale image of size 32x32 pixels. The first layer performs a convolution on the input image, followed by a subsampling operation. This process is repeated in the second layer, resulting in a reduced feature map. For classification, the final two layers that are totally connected are used. LeNet-5 was trained on the MNIST database, which contains 60,000 training images and 10,000 test images of handwritten digits, and achieved an error rate of less than 1% on the test set. The success of LeNet-5 demonstrated the potential of CNNs for image recognition tasks and paved the way for the development of more complex deep learning models.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

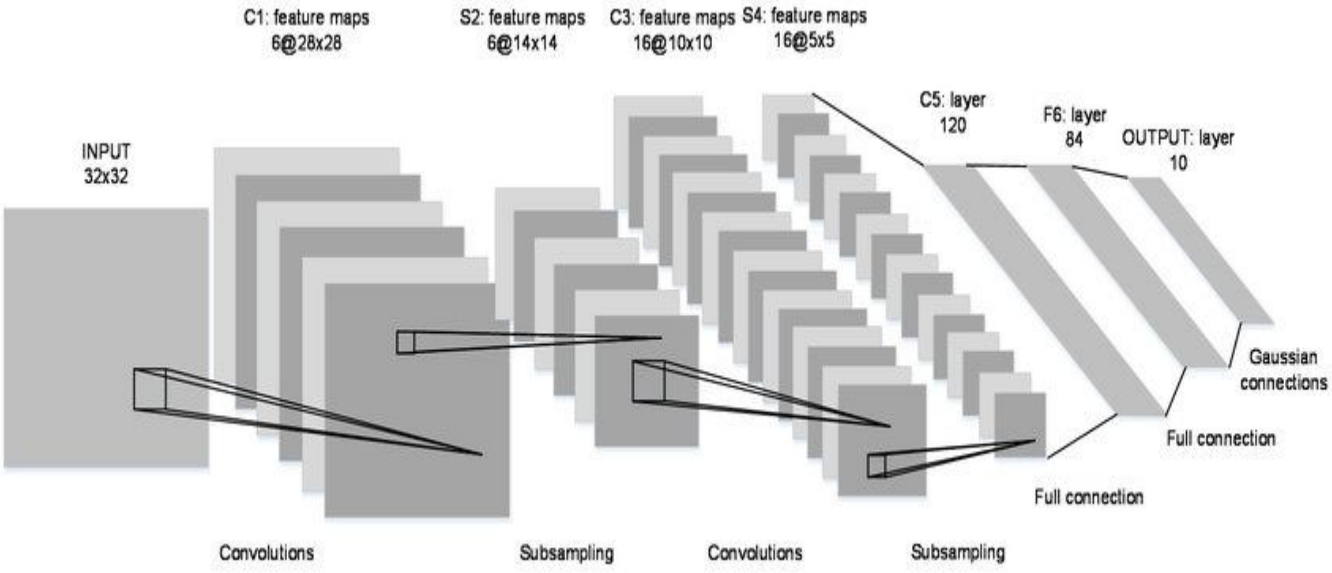


Fig 3.LeNet Architecture

D. Classification

Plant disease detection uses feature categorization. The plant's sick traits are eliminated, and the healthy leaf image is used to classify the plant. When the leaf is healthy and there is no classification, the findings are displayed as healthy; however, when there is a disease, which is indicated by black dots on a greyscale, the results are classified and the disease along with the classification's level of confidence are displayed. Two numerical arrays are used to classify data. Depending on the dataset provided, it is either a healthy leaf or a diseased leaf if the numerical arrays match. Classification is a quick yet important process that produces accurate results and is used to identify plant diseases.

E. Feature Extraction

Feature extraction is a process of extracting the most relevant and useful information or features from raw data in a way that the processed data can be used to train machine learning models or perform other analytical tasks.

In the context of machine learning, feature extraction is used to identify and extract important features from the raw input data that can be used to train a model. The goal is to reduce the dimensionality of the data while retaining the most informative features that are relevant to the task at hand. This can be achieved through techniques like principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA).

Feature extraction can also be used in other fields, such as computer vision and natural language processing. In computer vision, feature extraction involves analyzing images to identify relevant features such as edges, corners, or textures. In natural language processing, feature extraction involves transforming text data into a numerical representation, such as a bag-of-words model, that can be used to train a model.

The extracted features can then be used to train a machine learning model, such as a classification or regression model, to make predictions on new data. Feature extraction is an important step in the machine learning pipeline as it helps to reduce the dimensionality of the data, remove irrelevant information, and improve the accuracy and efficiency of the model.

1) *Max Pooling:* Maxpooling is a common operation used in convolutional neural networks (CNNs) for feature extraction from images. It is a form of subsampling that reduces the size of the feature maps produced by the convolutional layers. The operation works by dividing the input feature map into non-overlapping rectangular regions, called pooling windows or filters. For each pooling window, the maximum value within the window is extracted and used as the output value for that region. The result is a new, smaller feature map with reduced spatial dimensions but with preserved dominant features.

   Maxpooling has several benefits, such as: Translation invariance: Maxpooling reduces the sensitivity of the network to small translations of the input image. This means that the network can recognize the same object regardless of its position in the image. Dimensionality reduction: Maxpooling reduces the number of parameters in the network and helps prevent overfitting.

Robustness to noise: Maxpooling can help remove noise or small distortions in the input image by extracting the maximum value in the pooling window.

Typical maxpooling layers use a pooling window size of 2x2 with a stride of 2, resulting in a reduction of the feature map size by a factor of 2. However, other pooling window sizes and strides can be used as well depending on the network architecture and task requirements.

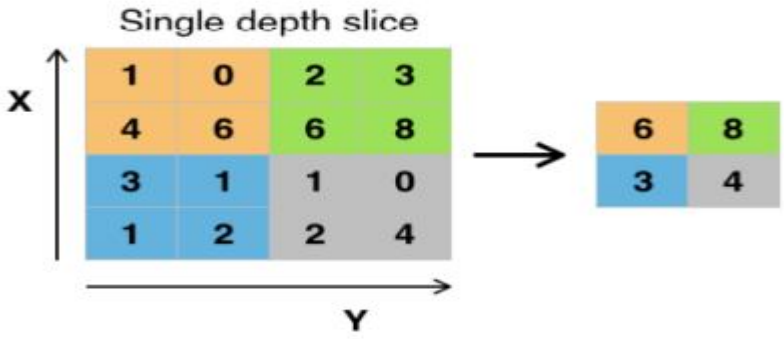


Fig 4: Pooling

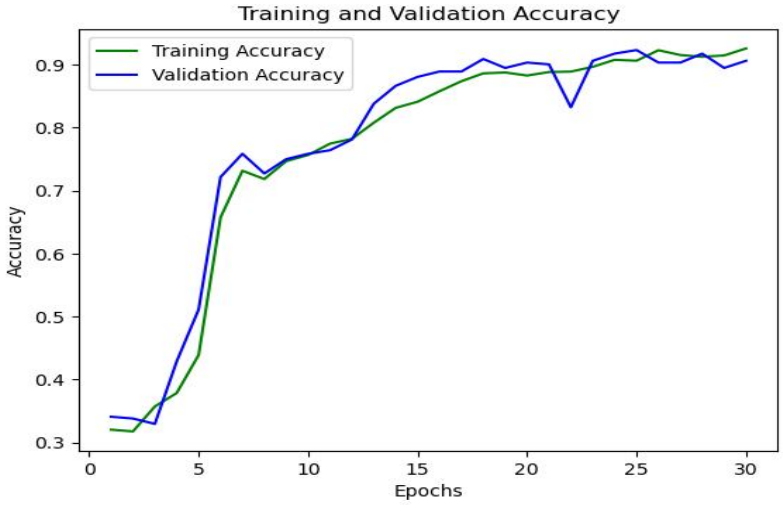


Fig 5. Training Performance of our model

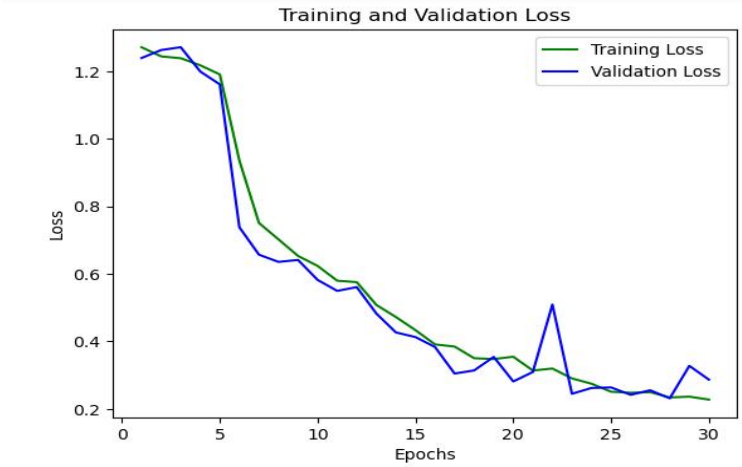


Fig 6. Learning curve of our model

IV. RESULTS AND DISCUSSION

A. *Results of the Training Model*

Our model's performance was assessed using a confusion matrix. We calculated precision and accuracy. The Dataset gave an accurate prediction of 0.95 and a precision of 0.90. This way our model predicted and operated well.

The precision that was attained by running a particular collection of epochs at every interval. It has been found that accuracy increases as the number of running epochs rises from interval to interval. Lastly, it is reported that the accuracy gained is 95% after executing 30 epochs at a time.

```
Epoch 27: val_loss did not improve from 0.24175
91/91 [==============================] - 433s 5s/step - loss: 0.2496 - accuracy: 0.9154 - val_loss: 0.2552 - val_accuracy: 0.
9034
Epoch 28/30
91/91 [==============================] - ETA: 0s - loss: 0.2339 - accuracy: 0.9126
Epoch 28: val_loss improved from 0.24175 to 0.23135, saving model to D:\major_proj\Foliar-diseases-in-Apple-Trees-Prediction-
master\models\apple-224.h5
91/91 [==============================] - 438s 5s/step - loss: 0.2339 - accuracy: 0.9126 - val_loss: 0.2313 - val_accuracy: 0.
9176
Epoch 29/30
91/91 [==============================] - ETA: 0s - loss: 0.2363 - accuracy: 0.9147
Epoch 29: val_loss did not improve from 0.23135
91/91 [==============================] - 429s 5s/step - loss: 0.2363 - accuracy: 0.9147 - val_loss: 0.3275 - val_accuracy: 0.
8949
Epoch 30/30
91/91 [==============================] - ETA: 0s - loss: 0.2276 - accuracy: 0.9258
Epoch 30: val_loss did not improve from 0.23135
91/91 [==============================] - 428s 5s/step - loss: 0.2276 - accuracy: 0.9258 - val_loss: 0.2867 - val_accuracy: 0.
9062
```

Fig 7: Accuracy at 30/30 epochs

B. *Output*

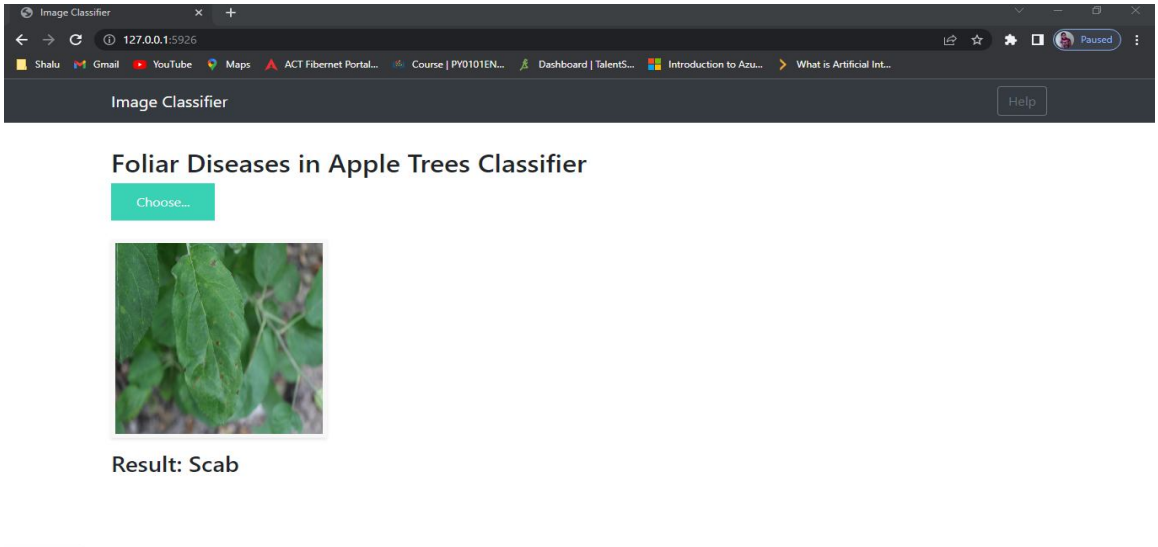


Fig 8: Output displaying over website

Above fig.7 shows, the output displayed over a website. After the successful completion of training, we uploaded an image as input and checked its activity. As a result, we obtained good accuracy in predicting disease**.**

V. ACKNOWLEDGMENT

We would like to thank our guide “Mr.V.N.L.N.Murthy” for helping throughout our project. Moreover, a particular thank you to the administration of Vardhaman College of Engineering for supporting us in completing the project successfully.

VI. CONCLUSION

Smallholder farmers rely on early and precise identification of crop diseases to avoid losses. A pre-trained convolutional neural network for apple leaf disease was deployed, and after that online deployment of a website for detecting plant diseases was the end outcome. All you need to utilize this service through a smartphone and an internet connection, and it is free. Overall, this study provides convincing evidence for the potential use of CNNs to support smallholder farmers in their fight against plant disease.

The implementation of a website was done for detecting the apple leaf disease detection using a convolutional neural network, so the Implementation of the project work was done successfully with an accuracy of 0.95 and we got to learn many more things from this project.

VII. FUTURE SCOPE

Plant leaf disease prediction has a significant future scope, especially with the advancements in machine learning and artificial intelligence. Here are some potential areas of growth for this field:

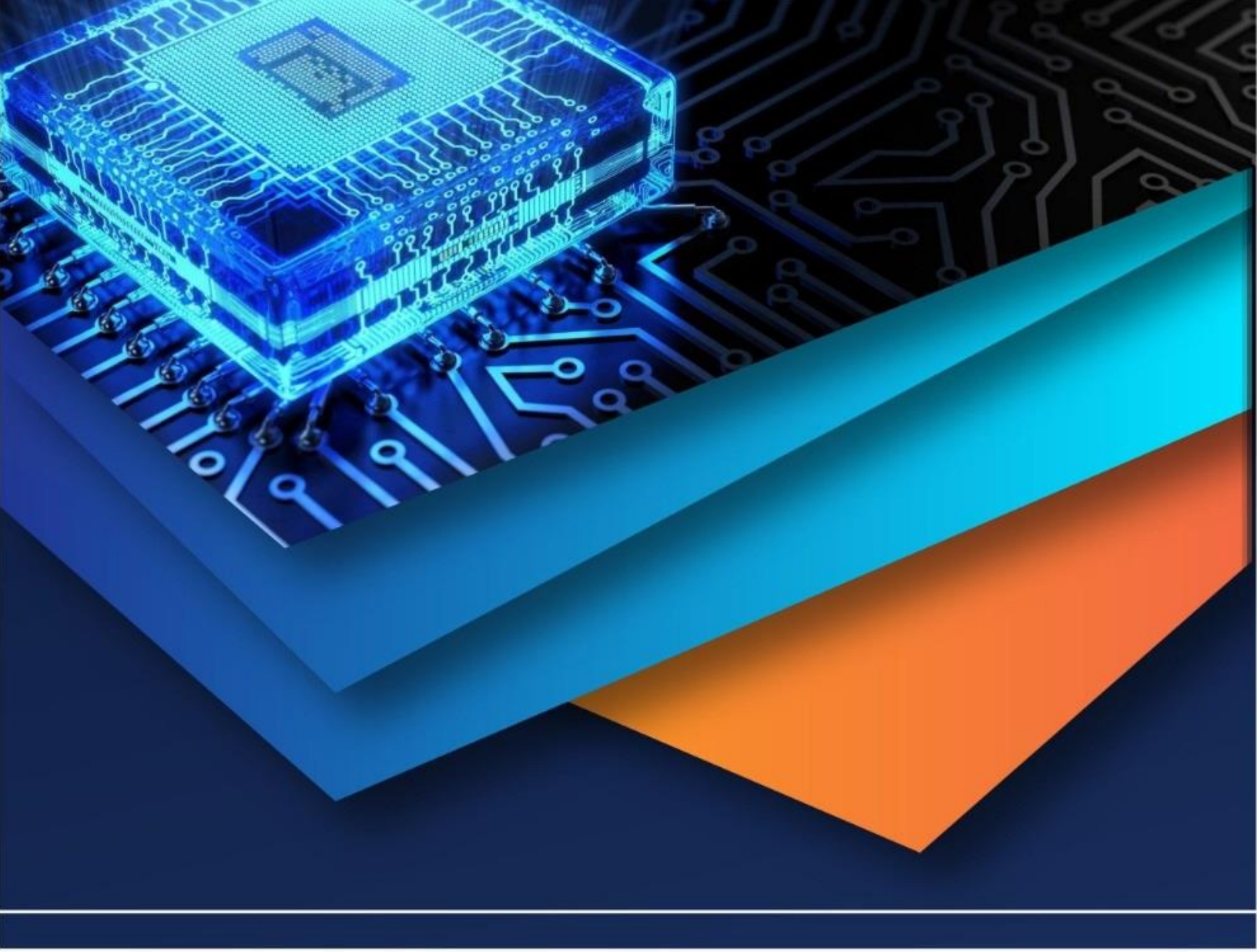
1) *Improved Accuracy:* With the development of more accurate machine learning algorithms, the accuracy of plant leaf disease prediction is likely to increase, making it easier for farmers to diagnose and treat plant diseases.
2) *Increased Efficiency:* The use of artificial intelligence can make the diagnosis of plant diseases more efficient, reducing the time and effort required to identify and treat diseases.
3) *Automation:* The automation of plant leaf disease prediction can significantly reduce labor costs associated with manual disease detection and provide a more accurate and comprehensive diagnosis of the plant's health.
4) *Precision Agriculture:* Plant leaf disease prediction can help farmers to implement precision agriculture practices, which can optimize crop yields and reduce waste, resulting in more sustainable and profitable farming practices.
5) *Global Impact:* Plant leaf disease prediction can have a significant global impact by reducing crop losses and increasing food security, particularly in developing countries where agriculture is a significant source of income and food production.

Overall, the future of plant leaf disease prediction looks promising, and continued advancements in technology are likely to further improve this field's accuracy and efficiency.

REFERENCES

[1] Liu, F., Wang, X., Fan, J., Cao, Z., & Zhou, T. Zhang, C. (2021). Review of deep learning models for leaf image-based autonomous disease diagnosis in plants. Agricultural Electronics and Computers, 182, 106008.
[2] Khan, M. A., Dauda, L., Al-Turjman, F., & Amin, M. B. (2020). A review of deep learning techniques for plant disease detection from leaf images. Applied Sciences, 10(5), 1615.
[3] Cruz, A. C. S., De Albuquerque, V. H. C., & De Araújo, A. F. (2018). Leaf disease detection using convolutional neural networks and transfer learning. Computers and Electronics in Agriculture, 153, 46-58.
[4] Zhang, Y., Zhang, J., & Li, X. (2021). A review of machine learning and deep learning methods for plant disease detection and classification. Computers and Electronics in Agriculture, 185, 106086.
[5] The 11th National Conference on Science and Engineering, Yangon, Yangon, Myanmar, Ko Ko Zaw, Zin Ma Ma Myo, and Wah Wah Hlaing. The detection and classification of leaf diseases using multiclass SVMs.
[6] Plant Disease Detection Using CNN, G. Shrestha, Deepsikha, M. Das, and N. Dey, sApplied Signal Processing Conference (ASPCON)IEEE 2020, pp. 109–113, 10.1109/ASPCON49795.2020.9276722.
[7] Khan, M. A., Dauda, L., Al-Turjman, F., & Amin, M. B. (2020). A review of deep learning techniques for plant disease detection from leaf images. Applied Sciences, 10(5), 1615.
[8] S. Sarkar, B. Ganapathysubramanian, and A. Singh (2016). High-throughput plant stress phenotyping using machine learning. 21(2), 110–124, Trends in Plant Science.
[9] S. P. Mohanty, D. P. Hughes, and M. Salathé (2016). using deep learning to identify plant diseases from images. Plant Science Frontiers, 7, 1419.
[10] Singh, D., Singh, P., & Kumar, A. (2021). Plant disease detection and classification using deep learning techniques: A review. Computers and Electronics in Agriculture, 184, 106012.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)