FarmHelp : Neural Network Based Plant Disease Detection System

Boora Satishchandra
Department of Information Technology
Terna Engineering College
Nerul(M.H), India 400706
satishboora47@gmail.com

Sah Smitasomya
Department of Information Technology
Terna Engineering College
Nerul(M.H), India 400706
smitasomyasah@gmail.com

Hrishikesh Mandhare
Department of Information Technology
Terna Engineering College
Nerul(M.H), India 400706
smitasomyasah@gmail.com

Dr. Vaishali Khairnar
HOD IT Department
Terna Engineering College
Nerul(M.H), India 400706
satishboora47@gmail.com

Prof. Smita Deshmukh
Department of Information Technology
Terna Engineering College
Nerul(M.H), India 400706
satishboora47@gmail.com

Prof. Vijayalaxmi Kadroli
Department of Information Technology
Terna Engineering College
Nerul(M.H), India 400706
satishboora47@gmail.com

***Abstract*—In growing nations like India, agriculture plays a huge role, but there are still major concerns about food security. Plant diseases are the primary cause of squandered crops. Early detection of plant diseases is crucial since they have an impact on how much food is produced. To recognize and categorize plant leaf diseases, a variety of machine learning models are employed. However, deep learning has a significant deal of potential for improved accuracy. This article provides a thorough explanation of how Deep Learning models are used to visualize different plant illnesses and offer appropriate advice for curing these problems. Our goal is to determine whether the plant is sick or not and if it is, to identify the type of illness that caused it. Along with this, the system will give a general overview of the illness, suggest a few straightforward home cures (if necessary), and direct the user toward getting the appropriate consultation. Convolutional Neural Network (CNN), a deep learning model is employed to recognize and categorize plant leaf diseases, to display the type of ailment as well as potential treatments after receiving input from the user in the form of picture data. Other programs utilized include Google Cloud/ AWS Cloud Server, and also TensorFlow which will be used for data preprocessing. Applying CNN algorithm, it is investigated that an average classification accuracy of 97.87% can be achieved to perform the classification of the leaf diseases. Hence, this would aid in increasing productivity and reducing crop loss.**

***Index Terms*—Plant Disease, TensorFlow, Convolution Neural Network (CNN), Deep Learning, Agriculture**

I. Introduction

Farming produces majority of the world's food and textiles. It can be called the science and art of growing crops, cultivating soil and raising livestock. Agriculture progressed slowly for about thousands of years. Farmers traditionally have used a variety of methodologies to protect their crops from pests and diseases. To control insects, they have used herb-based venoms on crops, hand-picked insects from the plants, bred strong crop varieties, and rotated crops. Pest control chemicals are now used by nearly all farmers, particularly in developed countries. Crop losses and prices have dropped dramatically as a result of the use of chemicals. Most crops are lost due to a lack of storage space, transportation, and plant diseases. More than 15 percent of crops in India are lost due to disease, making it one of the most pressing issues to address. The ability to produce enough food to meet the needs of more than 7 billion people is still preserved by modern technology. However, a number of factors, such as temperature change, the decline in pollinators, plant diseases, and others continue to threaten food security[1]. In addition to posing a threat to global food security, plant diseases can have negative effects on small holder farmers whose livelihoods depend on robust yields. To increase the growth and productivity of agricultural plants, diseases must be identified and treated. For traditional machine vision-based plant diseases and pests detection methods, conventional image processing algorithms or manual design of features plus classifiers are often used[2]. The Deep Learning (DL) approach is a subcategory of Machine Learning (ML), which was introduced in 1943 when threshold logic was introduced to build a computer model closely resembling the biological pathways of humans. This field of research is still evolving; its evolution can be divided into two time periods i.e from 1943–2006 and from 2012–until now. Machine Learning (ML) models have been employed for the detection and classification of plant diseases but, after the advancements in Deep Learning (DL), this area of research appears to have great potential in terms of increased accuracy. Also, many developed/modified DL architectures are implemented along with several visualization techniques like box plots, histograms, charts, heat maps etc. to detect and classify the symptoms of plant diseases[4]. DL has been widely used in image and video processing, voice processing, and natural language processing. Along with this, it has also become a research hotspot in the field of agricultural plant protection, such as plant disease recognition and pest range assessment, etc. The application of deep learning in plant disease recognition can avoid the disadvantages caused by artificial selection of disease spot features, also make plant disease feature extraction more objective, and improve the efficiency of research and technology transformation speed[6].

II. Literature Review

The goal of the literature review is to understand the knowledge that is already available on the Plant Disease Classification System. The study of the research helped in choosing the best algorithm and feature extraction technique for effective results.

Marzougui, F., Elleuch, M., Kherallah, M. [1], used pre-trained weights as a starting point to avoid a very long treatment. Followed by this, the proposed approach was compared to several artisanal shallow structure approaches based on machine learning. The proposed system achieved promising precision results on the plant leaves dataset, demonstrating the effectiveness of its approach for the detection of diseases.

The review by Liu, J., Wang, X. [2] provided a definition of plant diseases and pest detection problems and put forward a comparison with traditional plant diseases and pest detection methods. According to the difference in network structure, this study outlined the research on plant diseases and pests detection based on deep learning in recent years from three aspects of classification network, detection network, and segmentation network, and the advantages and disadvantages of each method were summarized. Common datasets were introduced,and the performance of existing studies was compared. On this basis, possible challenges in practical applications of plant diseases and pest detection based on deep learning were discussed. In addition, many possible solutions and research ideas were proposed for the challenges, and several suggestions were given.

Pandian JA, Kumar VD, Geman O, Hnatiuc M, Arif M, Kanchanadevi K [3], proposed a fourteen-layered 'deep convolutional neural network' (14-DCNN) to detect diseases in plant leaf using leaf images. A new dataset was created using numerous open datasets. 'Data Augmentation' techniques were then used to balance out the individual class sizes of the dataset. Three image augmentation techniques were used namely basic image manipulation (BIM), deep convolutional generative adversarial network (DCGAN) and neural style transfer (NST). The dataset consisted of about 147,500 images of 58 different 'healthy' and 'diseased' plant leaf classes along with one 'no-leaf' class. The proposed D-CNN model was trained in multi-graphics processing units (MGPUs) environment for exactly 1000 epochs. The random search with the coarse-to-fine searching technique was used to select the best hyper-parameter values in order to improve the training performance of the proposed model.

Muhammad Hammad Saleem, Johan Potgieter and Khalid Mahmood Arif in [4] provided a comprehensive explanation of Deep Learning models used to visualize various plant diseases. In addition to this, some research gaps were identified from which to obtain greater transparency for detecting diseases in plants, even before their symptoms appear clearly. The paper also stated that plant diseases affect the growth of their respective species, thus making its early identification very important.Several Machine Learning models had been employed for the detection and classification of plant diseases but, after the advancements in a subdivision of ML i.e Deep Learning (DL), this area of research appeared to have huge potential in terms of increased accuracy. Several developed DL architectures were implemented along with several visualization techniques to detect and classify the ailments. Also, several performance metrics were used for the assessment of these architectures or techniques.

Reference [5] dealt with a replacement approach to the development of a disease recognition model, supported leaf image classification, by the utilization of deep convolutional networks. All the essential steps required for implementing this disease recognition model were completely described throughout the paper, starting from gathering images to make a database, assessed by agricultural experts, and a deep learning framework to perform the deep Convolutional Neural Network training. The advance and novelty of the developed model illustrated its simplicity: healthy leaves and background images are in line with other classes, enabling the model to distinguish between diseased leaves and healthy ones, or from the environment by using CNN.

Reference [6] provided the research progress of deep learning technology in the field of crop leaf disease identification in recent years. The application of deep learning in plant disease recognition can avoid the disadvantages caused by the artificial selection of disease spot features, make plant disease feature extraction more objective as well as improve the research efficiency and technology transformation speed.In this paper, the authors L. Li, S. Zhang and B. Wang, presented the current trends and challenges for the detection of plant leaf disease using deep learning (DL) and advanced imaging techniques with the hopes that this research will be a valuable asset for those researchers who study the detection of plant diseases and pests. At the same time, they also discussed some of the current challenges and hurdles that need to be resolved.

A. KP and J. Anitha in [7], trained a deep learning model i.e. Convolutional Neural Network (CNN) to classify the different plant diseases. This model was used due to its massive success in image-based classification. The deep learning model provided a faster and more accurate predictions as compared to manual observations of the plant leaf. Here, the CNN model and pre-trained models namely VGG, ResNet, and DenseNet were trained using the dataset. Among them, the DenseNet model was observed to have achieved the highest accuracy.

Reference [8] presented a comprehensive review of the literature which aimed to identify the state of the art of the use of Convolutional Neural Networks (CNNs) in the process of diagnosis and identification of plant pest and diseases. Additionally, it presented some issues that were facing the models performance, and also indicated gaps that should be addressed in the future. From this review, it is possible to understand the innovative trends regarding the use of CNN's algorithms in the plant diseases diagnosis and to identify the gaps that need the attention of the research community.

Sai Reddy, B., Neeraja, S. [9] presented an automatic plant leaf damage detection and disease identification system. The first stage of the proposed method identified the type of the disease based on the plant leaf image using a method known as DenseNet.Stage two identified the damage in the leaf using deep learning-based semantic segmentation. Each 'RGB' pixel value combination in the image was extracted followed by supervised training that was performed on the pixel values using the 1D Convolutional Neural Network (CNN). The third stage suggested a remedy for the disease based on the disease type and the damage state observed. This gave an accuracy of 97 %.

Geetharamani G., Arun Pandian J. [10] put forth a novel plant leaf disease identification model following 'deep convolutional neural network' (Deep CNN) concept. Six types of data augmentation methods were used namely: image flipping, gamma correction, noise injection, principal component analysis (PCA) colour augmentation, rotation, and scaling. Compared to popular transfer learning approaches, the proposed model achieved a better performance when using the validation data. The model achieved 96.46 % classification accuracy.

G. Shrestha, Deepsikha, M. Das and N. Dey in [11] proposed a 'CNN' based method for plant disease detection. Simulation study and analysis was done on dummy images in terms of 'time complexity' and the 'area of the infected region' using image processing technique. A total of 38 cases had been fed to the model, out of which 26 cases were of 'diseased' plant leaves . Test accuracy was 88.80 %. Different performance matrices were derived for the same.

M. Sardogan, A. Tuncer and Y. Ozen [12] presented a 'Convolutional Neural Network (CNN) model and Learning Vector Quantization (LVQ) algorithm' based methodology for tomato leaf disease detection and categorization. The dataset contained about 500 images of tomato leaves with four symptoms of diseases. CNN for automatic feature extraction and categorization was modelled and 'Color' information was actively used for plant leaf disease researches. The filters were applied to three channels based on RGB components. Lastly, the LVQ had been fed with the output feature vector of convolution part for training the network.

III. Problem Definition

Problem Identified:

India is renowned as a nation of farmers, and there are acres dedicated to agriculture. The plants become sick because of a variety of circumstances, including temperature, pollution, pests, and many more. Sometimes these infected plants are not visible enough to determine the type of illness that it has produced. This results in mass deduction of food production.

Solution Proposed:

As per the literature survey, various research papers state the use of deep learning methods to provide solution to the problem stated above, out of which CNN is highly used. While some models focus on specific plants, others pose as a general solution to various plants for disease identification. This project focuses on studying and identifying whether or not the plant is diseased. If so, we provide a brief overview of the illness, suggest various natural treatments, and advise the user to seek an appropriate medical advice. As for the working, the 'Convolutional Neural Network' deep learning approach utilized here receives input from the user and shows the type of disease, along with a description of it and possible treatments. The goal of this project is to pinpoint the disease's type and offer a cure. Consequently, it contributes to increased output and the prevention of crop loss to a greater extent.

IV. Proposed Methodology

A neural network used for deep learning has numerous layers. It is a subset of machine learning, and it also does away with some of the standard data pre-processing steps. These neural networks use a combination of data inputs, weights, and bias using a large quantity of data in an effort to simulate the human brain. Additionally, it has covert layers that could improve and optimize the accuracy. Similar to this, real-world deep learning systems are so seamlessly integrated into goods and services that customers are completely oblivious to the intricate data processing going on behind the scenes. One of the common libraries for implementing neural networks is Tensorflow. Adopting best practices for model tracking, performance monitoring, and data automation.

A. Image Acquisition

Image acquisition is the very first step in image processing. This step is also called as pre-processing in image processing. It includes retrieving the image from a source. In image acquisition, a plant image is acquired and used as the target image for further processing in the system. There are fourteen different kind of diseased plant leaves in this project namely apple, blueberry, cherry, corn, grape, orange, peach, potato, pepper, raspberry, soybean, squash, strawberry and tomato. This would make up a total of 12 classes to be classified including 'healthy' plant leaf as one of them. About 80% of the images of each class will be used for training the classifier while the remaining 20% will be used for testing and validation. In Fig. 2 we can see a sample of different leaf image from each class in the dataset.

B. Cleaning and Image restoration

The next step after acquiring plant leaf image is cleaning of data which includes in improving the quality of the raw image and enhance the features of the image so that more information can be extracted from it. There exist two image restoration steps which is resizing and contrast enhancement. All the images obtained from Plant Village dataset are to be resized to 256 x 256 to ensure linearity in the datasets. After resizing the images, the contrast of the image will then enhanced to make the objects in the images more distinguishable to be processed. Data augmentation will also be applied to the dataset images.

Here data augmentation has to be done as we may not have enough diverse sets of images and so we need to rotate, flip and adjust contrast to create more training sample

C. Segmentation

One of the most difficult steps of image processing is segmentation. It involves splitting an image into its constituent parts or objects. The main goal of image segmentation is to obtain the region of interest (ROI) of the diseased plant leaf from image. In this scenario, the background of the dataset, which is in grey includes no information while the foreground is the plant leaf, which will be analyzed by the system. Hence, a mask must be created to mask out the background of the image which will result the image to be left with only pixels containing the image of the leaf with black background. After masking the image background, the healthy part of the leaf image will be then masked for calculating the diseased affected area in percentage. Calculation of the affected is done by the following formula:

$$\text{Affected Area (\%)} = \frac{\text{Diseased Area}}{\text{Total area of the leaf}} \times 100\%$$

$$x_{ij}^{l} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \omega_{ab} y_{(i+a)(j+b)}^{l-1}$$

This is just a convolution, which we can express in 'Matlab' using

conv2(x, w, 'valid ')

Then, the convolutional layer applies its nonlinearity:

$$y_{ij}^{l} = \sigma(x_{ij}^{l})$$

D. Feature Extraction

Once the image segmentation is completed, the segmented image features would be extracted to train the classifier. The algorithm that we will be using is CNN to perform feature extraction. Convolutional Neural Network or CNN is mainly used for image/ object recognition and classification. It is a type of artificial neural network and is used to detect diseasein plant leaves. In CNN, convolution and pooling are the two major functions where pooling is used to reduce the size of an image, and convolution is used to detect edges of patterns in an image. Suppose that we have some N × N square neuron layer which is followed by our convolutional layer.

E. Classification

Later, the features that were extracted from the previous stage will fed to the SVM classifier as inputs and it maps these features into high-dimensional feature spaces using its kernel function. SVM is a type of kernel-based supervised machine learning algorithm that would deliver analysis of data for classification tasks. In this work, multi class SVM is applicable as the basic SVM which only works with two classes. The algorithm will attempt to search natural clustering of the labelled training data to groups and then map new unlabelled data into these newly formed groups.

For User Interface design we will be using HTML,CSS and JavaScript. Lastly, the website will undergo testing and deployment on the Flask platform. In the website architecture, firstly the user has to upload the captured image of plant leaf. Then the applied algorithm will extract features from the data and compare it with the data from the dataset. Lastly the model will identify whether the plant leaf is diseasedor healthy and if diseased then its type of disease will also be predicted with appropriate remedies. Following is the working architecture of the user interface of the website for predicting diseased plant leaves.

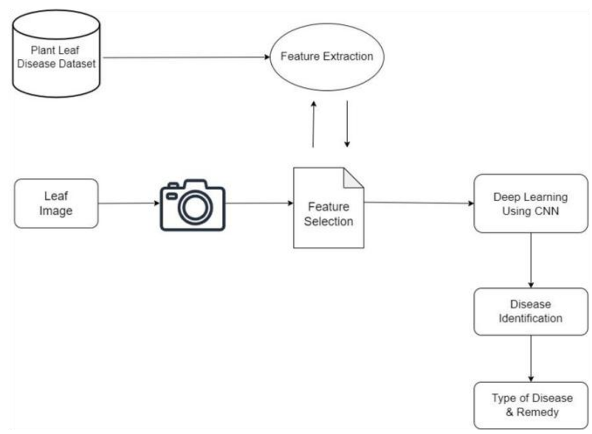

Fig. 1. Workflow of User Interface for plant disease prediction

Feature 1: The database consists of all the different plant leaf diseases which are taken into account as shown in the figure 1.

Feature 2: The module is trained repetitively to attain maximum accuracy. If a new image is given to the model then its features are compared with the features that will be already trained from the database. It will then provide the appropriate result.

Feature 3: The proposed system gives a gist of the disease detected along with some simple remedies (if applicable) to consider before consulting an expert.

V. Result

A database containing all the various plant leaf diseases that we will be considering can be seen in fig 2. To achieve the highest level of accuracy, the model is repeatedly trained. When the model receives a new image, its features will be compared to those that have already been trained and stored in the database. It will then deliver the necessary outcome.

To test the accuracy of this model we are using 10 % (20,638) images from the Plant Disease dataset and 10 % images for validation. These images are from 38 different classes. 10 % of each class is randomly selected for testing. Our model gives us more than 97.87 % accuracy on those images as well by telling whether the leaf is healthy or diseased.

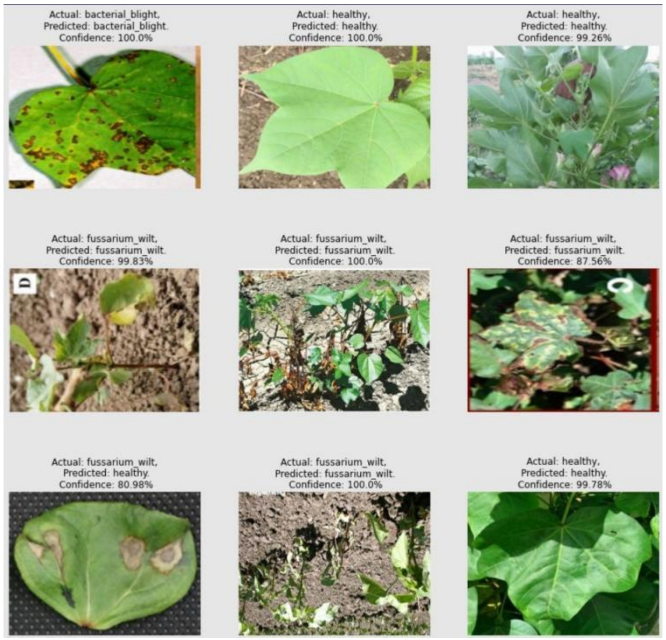

Fig. 2. Images of Database containing healthy and diseased leaves

Below is the Training and Validation accuracy and loss graphs generated by our model on testing dataset shown in figure 3. In accuracy graph displayed on left side below, the training accuracy is higher than the validation graph. As per the graph training accuracy has reached 97% and validation accuracy has reached 92% . Whereas in loss graph displayed on right side below, the training loss is lesser than the validation graph. According to the graph, training loss is 6% and validation loss is 29%. We have gained these accuracy and loss values because of training the model using 50 epochs to achieve such good result.

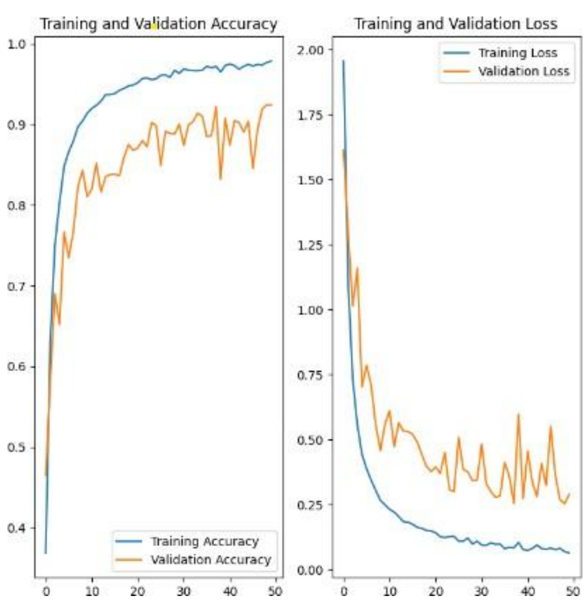

Fig. 3. Training and Validation graph of diseased plant leaves

Following is an image of our front-end wherein the user has to upload a captured image of a plant leaf which will help in identifying whether the plant leaf is healthy or diseased.

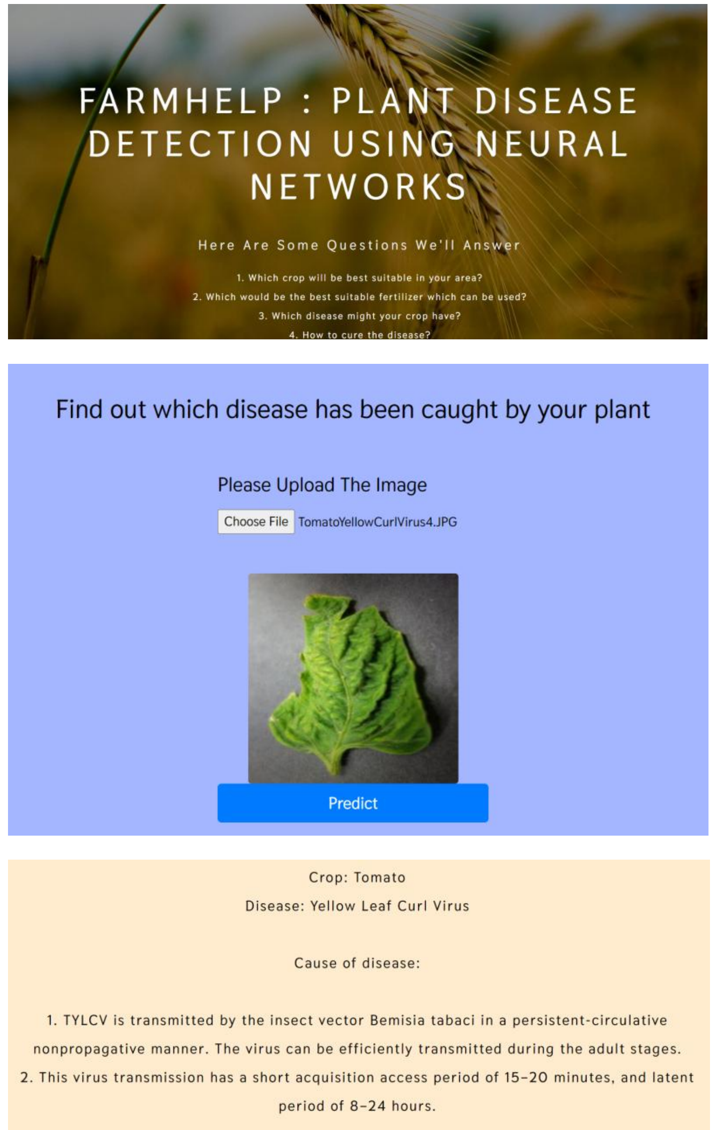

Fig. 4. User Interface for the user to upload an image

VI. Conclusion

In this paper, plant leaf disease detection and classification method are presented based on Convolutional Neural Network using Deep Learning. The dataset consists of 500 plant leaf images which consist of apple, blueberry, cherry,corn, grape, orange, peach, potato, pepper, raspberry, soybean, squash, strawberry and tomato leaves. Three different input matrices have been obtained for R, G, and B channels to start convolution on every image in the dataset. Each input image matrix of the dataset has been convoluted. Max pooling, reLU activation function, and softmax activation function have been implied to the output matrix. The analysis has been carried out on healthy and diseased leaf images to perform classification. It is concluded that the proposed method effectively recognizes different types of leaf diseases for the respective type of leaf. To improve recognition rate in the classification process different filters or different sizes of convolutions could be used as well.

VII. Future Scope

Future research will focus on real-time data gathering issues and multi-object deep learning models that can detect plant illnesses from a group of leaves rather than a single leaf. Additionally, we are attempting to integrate the trained model we developed into a mobile application. It will eventually assist farmers and the agricultural industry in identifying leaf diseases in real-time and offering advice or solutions.

References

[1] Marzougui, F., Elleuch, M., Kherallah, M. (2020). A Deep CNN Approach for Plant Disease Detection. 2020 21st International Arab Conference on Information Technology (ACIT). doi:10.1109/acit50332.2020.9300072

[2] Liu, J., Wang, X. Plant diseases and pests detection based on deep learning: a review. Plant Methods 17, 22 (2021). https://doi.org/10.1186/s13007-021-00722-9

[3] Pandian, J.A.; Kumar, V.D.; Geman, O.; Hnatiuc, M.; Arif, M.; Kanchanadevi, K. Plant Disease Detection Using Deep Convolutional Neural Network. Appl. Sci. 2022, 12, 6982. https://doi.org/10.3390/app12146982

[4] Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant Disease Detection and Classification by Deep Learning. PlanNetworks, 8, 468. https://doi.org/10.3390/plants8110468

[5] Srivastava, Prakanshu / Mishra, Kritika / Awasthi, Vibhav / Sahu, Vivek / Pal, Pawan Kumar. (2021). PLANT DISEASE DETECTION USING CONVOLUTIONAL NEURAL NETWORK. International Journal ofAdvanced Research. 09.691-698.10.21474/IJAR01/12346.

[6] L. Li, S. Zhang and B. Wan,g, "Plant Disease Detection and Chasification by Deep Learning—A Review," in IEEE Access, vol. 9, pp. 56683-56698, 2021, doi: 10.1109/ACCESS.2021.3069646

[7] A. KP and J. Anitha, "Plant disease classification using deep learning," 2021 3rd International Conferethe nce on Signtypesrocessing and Com- munithe cation (ICPSC), Coimthe batore, India, 2021, pp. 407-411, doi: 10.1109/ICSPC51sizes2021.9451696.

[8] S. M. Omer, K. Z. Ghafoor, and S. K. Askar, "Plant Disease Diagnosing Based on Deep Learning Techniques: A Survey and Research Challenges", ARO, vol. 11, no. 1, pp. 38-47, Feb. 2023.

[9] Sai Reddy, B., Neeraja, S. Plant leaf disease classification and damage detection system using deep learning models. Multimed Tools Appl 81, 24021–24040 (2022). https://doi.org/10.1007/s11042-022-12147-0

[10] Geetharamani G., Arun Pandian J., Identification of plant leaf diseases using a nine-layer deep convolutional neural network, Computers Electrical Engineering, Volume 76,2019,Pages 323-338,ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2019.04.011. (https://www.sciencedirect.com/science/article/pii/S0045790619300023)

[11] G. Shrestha, Deepsikha, M. Das and N. Dey, "Plant Disease Detec- tion Using CNN," 2020 IEEE Applied Signal Processing Conference (ASPCON), Kolkata, India, 2020, pp. 109-113, doi: 10.1109/ASP- CON49795.2020.9276722.

[12] M. Sardogan, A. Tuncer and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm," 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia and Herzegovina, 2018, pp. 382-385, doi: 10.1109/UBMK.2018.8566635