International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 03 | Mar 2023 www.irjet.net p-ISSN: 2395-0072

A Review: Plant leaf Disease Detection Using Convolution Neural Network in Machine Learning

Akshay Misal[1], Sagar Golhar[1], Triveni Awachat[1], Mohammad Muzaif[1], Mrs. Bharati Chamat[2]

[1]Student, Dept of Information Technology, KITS, RAMTEK
[2]Assistant Professor, Dept. of Information Technology, KITS, RAMTEK

---***---

Abstract - *This paper provides an overview of the use of deep learning in plant protection, specifically in the identification of crop leaf diseases. Deep learning has gained significant attention due to its advantages in feature extraction and machine learning, making it a popular tool in various fields such as image and video processing, speech processing, and natural language processing. By applying deep learning to plant disease identification, it can provide more objective and efficient extraction of disease characteristics and improve research efficiency and technology transformation. our research aims to summarize recent advances in deep learning research related to crop leaf disease identification, highlighting current trends and challenges in this area, and serving as a useful resource for researchers studying plant pest detection. We are using CNN algorithm which provided 97% accuracy to detect disease.*

***Key Words* : CNN, feature Extraction, plant leaf disease detection. Accuracy.**

1.INTRODUCTION

Plants play a crucial role in providing energy and are a key component in addressing the global warming crisis. However, several plant diseases have the potential to cause disastrous economic, social, and ecological consequences. To meet the growing demand for food, global crop production must increase by at least 50% by 2050.

Currently, the majority of crop production occurs in Africa and Asia, where most growers are family-run businesses with limited horticultural experience. Consequently, pests and diseases often cause yield losses of over 50%. The traditional method of human analysis through visual inspection is no longer sufficient for identifying agricultural diseases.

The primary objective of this project is to detect the most common diseases in plant leaves, specifically those of tomato, potato, and pepper. This system can identify 15 different forms of diseases in these three plants. The user can upload an image of a leaf, and if the leaf is infected, the system will display the name of the disease and suggest the appropriate pesticide after clicking the predict button. If there is no disease present on the plant leaf, the system will show a message that reads "There is no disease on the plant." Additionally, the system displays the percentage of affected areas and recommends pesticides based on that proportion.

1.1.Need of Plant Detection

Plant diseases can have a devastating impact on both farmers and consumers. When plants become infected, farmers may experience a significant loss in income, while consumers may face shortages and higher prices. The economic cost of plant leaf diseases alone is estimated at a staggering 60 billion dollars.

One of the most critical requirements for this project is that the detection of disease must be easy for farmers to perform in a simple manner. Farmers cannot rely on visual inspection alone to predict the onset of disease in plants. Therefore, the proposed system must enable farmers to identify the disease and recommend effective pesticides to prevent further spread of the disease. If a single plant leaf becomes infected, it can quickly spread to other leaves and potentially destroy the entire yield.

1.2.Problem Statement

To optimize the use of Machine Learning algorithms and streamline the process of detecting plant diseases, as well as assessing their impact on crop yield and recommending suitable pesticides, with the aim of reducing both time and cost for farmers.

1.3.Scope

This project has the potential to significantly benefit rural farmers by assisting them in protecting their crops from diseases. Farmers often suffer significant losses due to crop diseases, and this system aims to prevent such situations from occurring. Moreover, the project has been implemented in a regional language to enhance its accessibility and comprehension for farmers.

Additionally, the system can recommend effective pesticide treatments and appropriate dosages based on the current disease situation. It can also provide farmers with insights into the impact of crop diseases on crop yield, allowing them to make informed decisions.

Overall, this project has the potential to be a valuable tool for farmers in combating crop diseases and ensuring the health and productivity of their crops.

1.4 Convolutional Neural Network

Convolutional neural networks (CNNs) may seem like an unconventional blend of biology and math with a touch of computer science, but they have revolutionized computer vision and become one of the most influential advancements in this field. In 2012, Alex Krizhevsky employed neural nets to win the annual ImageNet competition, reducing the classification error record from 26% to 15%, a remarkable feat at the time. Since then, numerous companies have integrated deep learning at the heart of their services. Facebook uses neural nets for automatic tagging algorithms, Google for photo search, Amazon for product recommendations, Pinterest for home feed personalization, and Instagram for search infrastructure.

Despite their versatility, CNNs are widely recognized for their effectiveness in image processing. In the realm of image processing, let's examine how CNNs can be utilized for image classification.

1.5 Modules

Feature Selection:- An important step in the process is image annotation, the goal of which is to label the locations and types of object spots in the infected photographs. A convolution neural networks (CNN) algorithm with a frame selection function is created in Python specifically for this stage.

Region Based Convolutional Neural Network:- A set of machine learning models called R-CNNs (Region-based Convolutional Neural Networks) are utilised in computer vision and image processing. Every R-primary CNN's objective is to identify objects in any input image and define borders around them because it was specifically created for object detection. The R-CNN model uses a process called selective search to extract details about the region of interest from an input image. The bounds of the rectangles can be used to depict a region of interest. There can be more than 2000 regions of interest, depending on the scenario. To create output features, CNN uses this area of interest. When working with photos, the data is typically enormous, and if it is fed to the model in its current form, the model will become cumbersome, the training process will take a long time, and the memory requirements will be high. Thus, a region convolutional neural network will be used. The Convolutional neural network's input layer receives the pre-processed data as input. A filter is used to help reduce the dimensions of the input data when reading the pre-processed data.

Classification:- We can see that the plant leaf has been impacted by the disease. Finally, the disease will be tagged with the types of diseases it affects and its treatment using the deep learning approach.

2. LITERATURE SURVEY

In this [1] paper The use of smart farming systems that incorporate advanced technologies such as deep learning and computer vision can greatly improve the efficiency and productivity of agriculture, including the cultivation of tomatoes. Tomato farming involves various factors such as soil quality, environmental conditions, and sunlight exposure, which can make it challenging to avoid diseases. However, the development of an innovative solution that utilizes an automated image capturing system can help detect and recognize leaf diseases in tomato plants.

The system uses a motor-controlled image capturing box to capture images of the four sides of every tomato plant, which are then analyzed using a deep convolution neural network (CNN) to identify the presence of three diseases, namely Phroma Rot, Leaf Miner, and Target Spot. To train the CNN, a dataset containing images of both diseased and healthy tomato leaves was collected. The system achieves a high accuracy of 95.75% using Transfer Learning disease recognition model.

The F-RCNN trained anomaly detection model produced a confidence score of 80%, which indicates the level of certainty in identifying the presence of a disease. The automated image capturing system was implemented in real-world settings and was able to achieve an accuracy rate of 91.67% in identifying tomato plant leaf diseases.Overall, the use of smart farming systems and advanced technologies such as deep learning and computer vision can greatly improve the efficiency and productivity of tomato farming by enabling the early detection and prevention of diseases, which can help increase the quality and quantity of tomato production.

In This [2] paper The field of agriculture has a great impact on our lives. Agriculture is the most important sector of our economy. It is difficult for farmers to identify leaf diseases, so their yields are lower. However, videos and leaf images give agronomists a better view and can provide better solutions. Therefore, problems related to crop diseases can be solved [2]. It is important to note that if the productivity of the crop is poor, there is a high chance that it will not be able to provide good nutrition [2]. Thanks to improvements and developments in technology, the devices are smart enough to identify and detect plant diseases. Recognize diseases and treat them more quickly to reduce the negative impact on the harvest [2]. This article mainly investigates the use of image processing technology for the detection of plant diseases [2]. This paper obtains an open dataset image composed of 5000 images of healthy and diseased plant leaves, and uses semi-supervised techniques to detect crop types and four disease types [2].

Scientists used the convolutional neural network (CNN) to divide plant leaf diseases into 15 categories, including three classes for healthy leaves and 12 classes for diseases found in various plants, such as bacteria, fungi, and others. With a training accuracy of 98.29% and a testing accuracy of 98.029% for all used data sets, they were able to achieve excellent accuracy in both training and testing. [3]

An overview of feature extraction with GLCM, HSV-dependent classification, and picture segmentation with K-means clustering for locating infected leaf regions. The proposed methodology can reliably identify and classify plant diseases with 98% accuracy when processed by a Random Forest classifier. [4]

Researchers developed an integrated deep learning framework that uses a pre-trained VGG-19 model for feature extraction, and a stacking ensemble model to detect and classify leaf diseases from photographs, in order to reduce productivity and financial losses in the agriculture industry. A dataset containing 3242 pictures total and two classifications (Infected and Healthy) was used to test the system. Their efforts have been compared to other contemporary algorithms (kNN, SVM, RF and Tree). [5].

For automatic feature extraction and categorization, a CNN was suggested. The study of plant leaf diseases makes considerable use of colour data. Three channels receive filtering based on RGB model components. The output feature vector from the convolution component was fed into the LVQ for network training [6].

The main objective was to decrease the usage of pesticides in order to increase crop yield and productivity. Plant disease detection can be done via image processing. Some of the steps in the illness detection process include preprocessing the image, feature extraction, classification, and disease prediction. The examination of high-resolution photographs of the plant for appropriate therapy and prevention can therefore be aided by the development of a recognition system [7].

Deep learning techniques were used to diagnose diseases. The most significant challenge in the implementation was selecting a deep learning architecture. As a consequence, AlexNet and SqueezeNet, two different deep learning network topologies, were tested. Both of these deep learning networks were trained and validated on the Nvidia Jetson TX1. Photos of tomato leaves from the Plant Village dataset were used for the training. There are ten different courses, and each one has positive images. Moreover, web images are utilised to test trained networks.[8]

Two different models, Faster R-CNN and Mask R-CNN, are used in these techniques in [9], with Faster R-CNNusing Mask R-CNN to locate and segment the shapes of the affected spots and being used to identify the different tomato diseases. To determine which deep convolutional neural network best fits the tomato illness detection task, four distinct deep convolutional neural networks are combined. Data for the dataset was obtained from the Internet and is broken down into three sets: a training set, a validation set, and a test set utilised in the experiments. The trials' findings showed that their suggested models are capable of accurately and swiftly identifying the eleven tomato disease subtypes as well as classifying the locations and shapes of affected patches. This system's primary goal is to properly identify tomato plant problems using IoT, Machine Learning, Cloud Computing, and Image Processing [10].

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

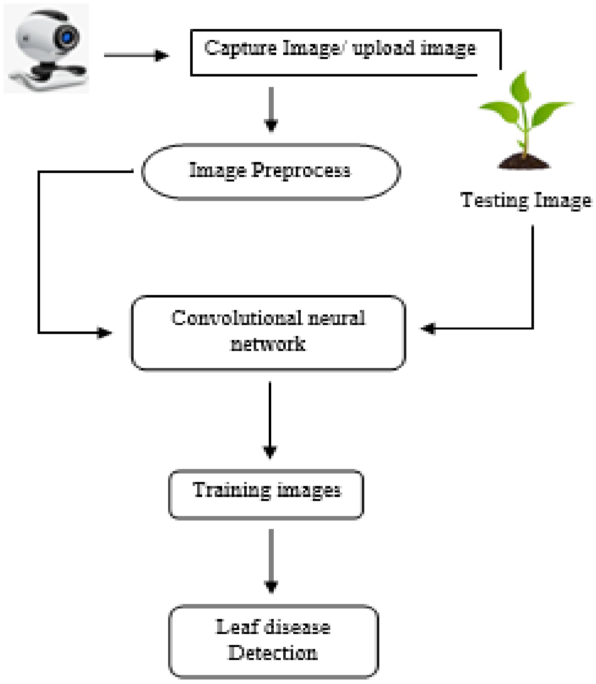
Now a day, we regularly use digital camera and other electronic devices day today life. Then the automatic plant disease identification has been widely applied as a satisfactory alternative. Most of the cases following traditional machine learning approaches such as support vector machine (SVM), CNN and K-means clustering have complex image pre-processing and feature extraction steps, which reduce the efficiency of disease diagnosis

3.2 PROPOSED SYSTEM

Agriculture is one of the most important occupations in the world. It plays an important role because food is a basic need of all living beings on this planet. In this proposed system, deep learning region-based convolutional neural network (R-CNN) method is used for recognition. They have two phases, the training phase and the testing phase. At the initial stage, they perform image acquisition, preprocess the images, and train the images using R-CNN. The second step is the classification and identification of leaf diseases. For training purposes, images are extracted from the dataset, while real-time images are available for testing. Foliar disease diagnosis is performed from images uploaded to the system or database. Real-time environment input requires image pre-processing and then performs feature classification to detect disease diagnosis and get disease name.

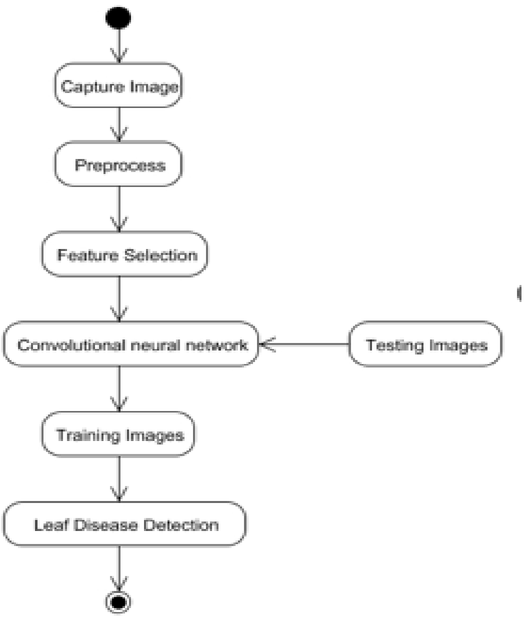
3.2.1 Architecture Diagram

In below figure this is architecture of our project. User can capture and upload the image that want to predicate. The train module which consists of different plant leaf disease that already train and store in database. If new image inserted that match with dataset and show appropriate result to the user



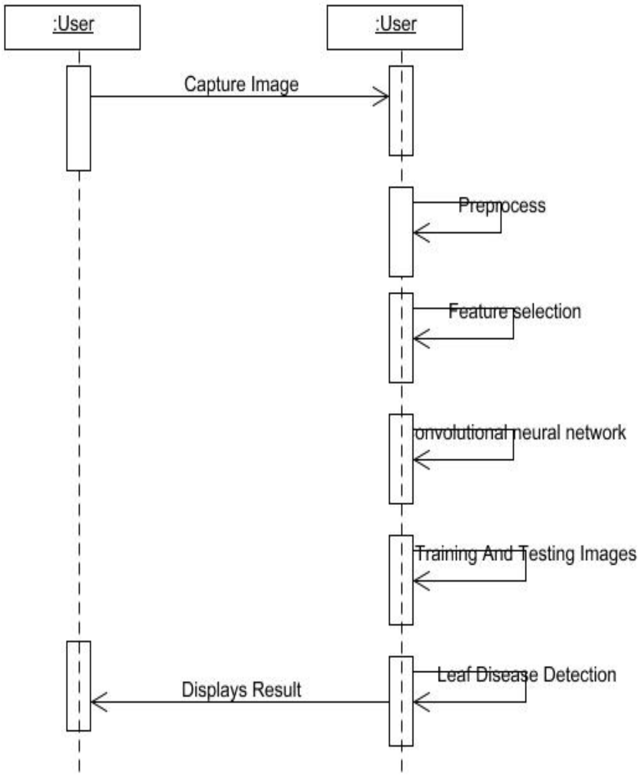
3.2.2 Activity Diagram

In below figure this is activity of software. The capture image gets preprocess and the feature selection is done. CNN module is used to map with train image and for classification on image



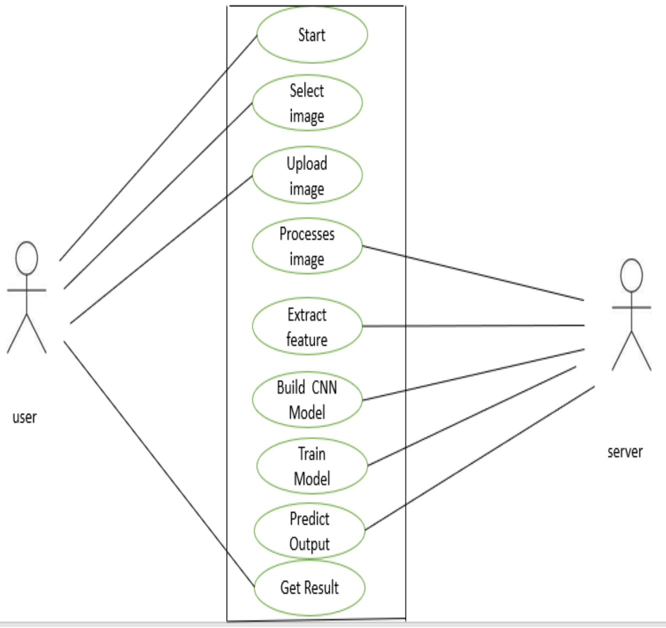
3.2.3 Sequence Diagram

The below is a sequence diagram is an interaction diagram that demonstrates the order in which processes collaborate with one another. It is made up of message sequence diagrams, also known as event diagrams, sequence diagrams, and message flow diagrams.



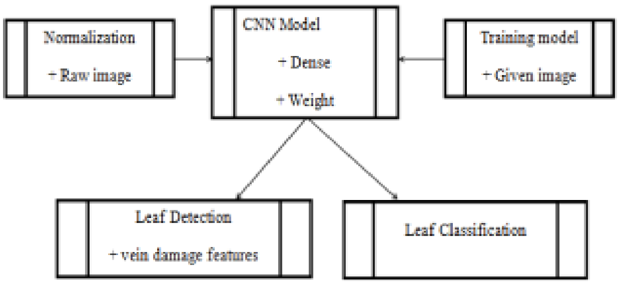
3.2.4 Use Case Diagram

The below is use case diagram of project the main component are user and server and their action can show in use case diagram,



3.2.4 Class Diagram

The below figure is class diagram. A CNN model composed of densities and weights is introduced into the normalized class of the original image. CNN models use trained models for classification and detection. The image dataset constitutes the training model class. These attributes are used for sheet detection.



4. Dataset

We download the dataset from plant village site which having 32 class with healthy and non-healthy. The dataset having 53,655 images fallowing are class we are using.

	no. of images
Apple___Apple_scab	2016
Apple___Black_rot	1987
Apple___Cedar_apple_rust	1097
Apple___healthy	2008
Blueberry___healthy	1816
Cherry_(including_sour)___healthy	674
Corn_(maize)___Common_rust_	1907
Corn_(maize)___healthy	1859
Corn_(maize)___Northern_Leaf_Blight	690
Grape___Black_rot	1888
Grape___Esca_(Black_Measles)	1109
Grape___healthy	1692
Peach___Bacterial_spot	1838
Peach___healthy	1728
Pepper,_bell___Bacterial_spot	822
Pepper,_bell___healthy	1988
Potato___Early_blight	1939
Potato___healthy	1824
Potato___Late_blight	1939
Raspberry___healthy	1781
Soybean___healthy	2022
Squash___Powdery_mildew	1731
Strawberry___healthy	1824
Strawberry___Leaf_scorch	1773
Tomato___Bacterial_spot	1702
Tomato___Early_blight	1920
Tomato___healthy	1925
Tomato___Late_blight	1835
Tomato___Leaf_Mold	1882
Tomato___Septoria_leaf_spot	1596

5. CNN MODEL STEPS

Conv2D: It is the activation layer that converts an image into multiple images is the activation function.

MaxPooling2D: This is used to maximize the value of the matrix of the given size, also used for the next 2 layers.

Flatten: allows you to flatten the size of the image obtained after folding. Dense: This is used to make this a fully connected model, which is a hidden layer. Drop: This is used to avoid overfitting on the dataset and is a dense output layer containing a single neuron that decides which class the image belongs to.

Image Data Generator: This is the function to resize image, apply offset in certain range, scale image and scroll image horizontally This image data generator includes all possible orientations of the image.

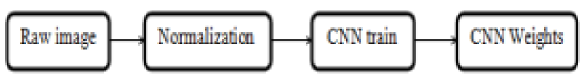
Training Process: Train datagen. Flow from directory is a function used to prepare data from the directory train_dataset Target_size specifies the target size of the image. Generation of test data. flow_from_directory is used to prepare test data for the model, everything is similar to the above. fit generator is used to fit the data in the model created above, other factors used are steps_per_epochs which tells us how many times the model will run for the training data.

Epochs: This tells us how many times the pattern will be trained in both forward and reverse passes.

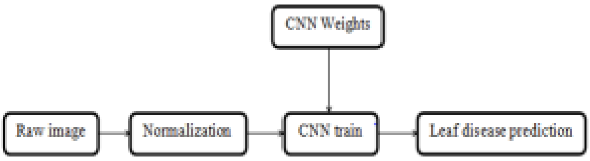
validation process: validation data is used to feed validation/test data into the model. validation steps indicate the number of validation/testing samples.

6. TRAINING AND TESTING MODEL

The dataset is preprocessed, including image scaling, image reshaping, and array form conversion. Similar processing is likewise done on the test image. Every image from a dataset of around 32 distinct plant leaf diseases can be used as a test image for the software.

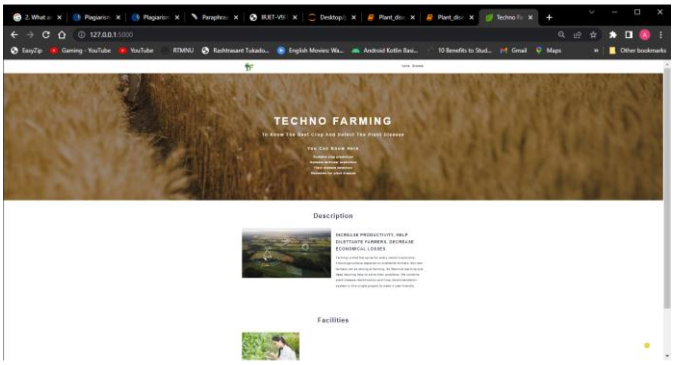


The model (CNN) is trained using the training dataset so that it can recognize the test image and the disease it possesses. Convolution2D, MaxPooling2D, Dropout, Activation, Flatten, and Dense layers are some of the layers that CNN has. If the plant species is present in the dataset after the model has been successfully trained, the programmed can detect the disease. In order to forecast the disease, the test image and trained model are compared after effective training and preprocessing.

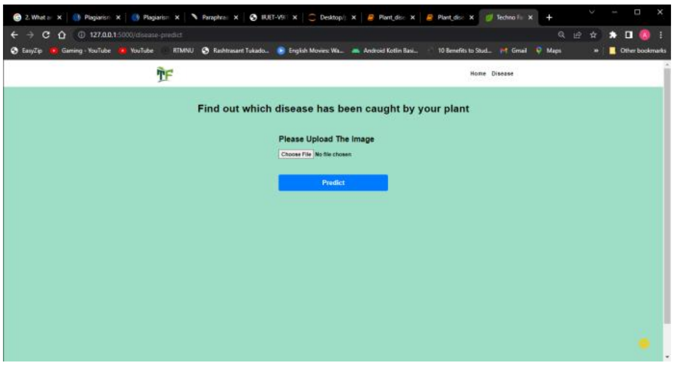


6.RESULT

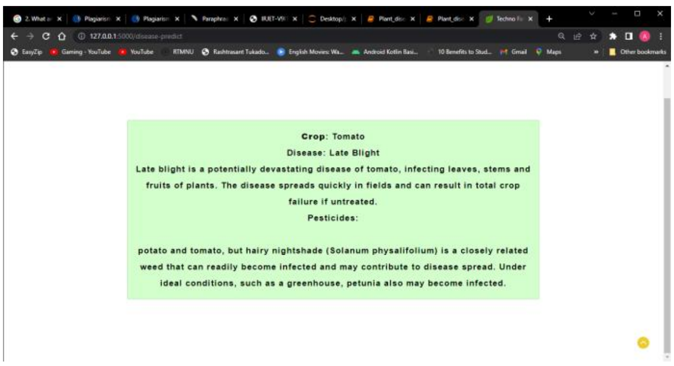
This is home screen of our project.



After click on disease the next screen will be open where we have to select the infected plant leaf to check whether is healthy or not



After click on predicate the result will be show.



7. CONCLUSIONS

We create web based application to ensure that give 96% accurate result.

It concentrated on applying a CNN model to forecast the pattern of plant diseases using images from a given dataset (a trained dataset) in the field and historical data. Some of the benefits of this following findings regarding plant leaf disease prediction. This method will cover the greatest variety of plant leaves, allowing farmers to learn about leaves that may have never been cultivated. By listing all potential plant leaves, it aids farmers in choosing which crop to grow. Additionally, this technology takes historical data production into account, giving the farmer information into market prices and demand for specific plants.

REFERENCES

[1] Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala, (2019)"Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition," International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)

[2] Suma VR Amog Shetty, Rishab F Tated, Sunku Rohan, Triveni S Pujar,(2019) "CNN based Leaf Disease Identification and Remedy Recommendation System," IEEE conference paper

[3] Marwan Adnan Jasim and Jamal Mustafa AL-Tuwaijari , "Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques", 2020 International Conference on Computer Science and Software Engineering, IEEE 2020.

[4] Poojan Panchal, Vignesh Charan Raman and Shamla Mantri , "Plant Diseases Detection and Classification using Machine Learning Models", IEEE 2019.

[5] Jiten Khurana and Anurag Sharma ," An Integrated Deep Learning Framework of Tomato Leaf Disease Detection", International Journal of Innovative Technology and Exploring Engineering (IJITEE),2019.

[6] Melike Sardogan, Adem Tuncer and Yunus Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LV Algorithm", IEEE 2018.

[7] Gaurav Langar, Purvi Jain and Nikhil, "Tomato Leaf Disease Detection using Artificial Intelligence and Machine Learning", International Journal of Advance Scientific Research and Engineering Trends, 2020.

[8] Halil Durmus and Murvet Kirci, "Disease Detection on the Leaves of the Tomato Plants by Using Deep Learning ",International Conference onAgro -Geo informatics,2018

[9] Minghe Sun and Jie xue, "Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and Object Detection Techniques", Research Article, 2019.

[10] Saiqa Khan and Meera Narvekar, "Disorder detection of tomato plant(solanum lycopersicum) using IoT and machine learning", Journal of Physics: Conference Series ,2019.