International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

PLANT DISEASE IDENTIFICATION

Aboli G. Dhote[1], Krunal G. Dhote[2], Achal V. Kale[3] , Gnyaneshwar B. Yadav[4], Dr. N. R. Chopde[5]

[1] Aboli G. Dhote, Department of Computer Science and Engineering, P. R. Pote Patil Institute of Engineering and Research, Amravati, India

[2] Krunal G. Dhote, Department of Computer Science and Engineering, P. R. Pote Patil Institute of Engineering and Research, Amravati, India

[3] Achal V. Kale, Department of Computer Science and Engineering, P. R. Pote Patil Institute of Engineering and Research,Amravati, India

[4] Gnyaneshwar B.Yadav, Department of Computer Science and Engineering, P. R. Pote Patil Institute of Engineering andResearch, Amravati, India

[5]Dr. N. R. Chopde, Department of Computer Science and Engineering, P. R. Pote Patil Institute of Engineering and Research,Amravati, India

***

ABSTRACT

The Plant Disease Identifier using Convolutional Neural Networks (CNN) is a computer vision project that aims to accurately identify and classify plant diseases from images of leaves. The project utilizes a deep learning approach with a CNN architecture to learn and extract meaningful features from the input images, which are then used to classify them into various disease categories. The model is trained on a large dataset of images of healthy and diseased plants and is capable of identifying multiple diseases present in the same leaf. The accuracy of the model is evaluated using various performance metrics, such as precision, recall, and F1 score. The project has the potential to aid in early detection and treatment of plant diseases, thereby minimizing crop damage and increasing agricultural productivity.

1. INTRODUCTION

Modern technologies have given human society the ability to produce enough food to meet the demand of more than 7 billion people. However, food security remains threatened by a number of factors including climate change (Tai et al., 2014),the decline in pollinators (Report of the Plenary of the Intergovernmental Science-Policy Platform on Biodiversity Ecosystem and Services on the work of its fourth session, 2016), plant diseases (Strange and Scott, 2005), and others. Plant diseases are not only a threat to food security at the global scale but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. In the yield loss of more than 50% due to pests and diseases are common (Harvey et al., 2014). Furthermore, the largest fraction of hungry people (50%) lives in smallholder farming households (Sanchez and Swaminathan, 2005), making smallholder farmers a group that's particularly vulnerable to pathogen-derived disruptions in food supply.

Plant diseases are one of major reasons behind the production and economic losses in agriculture. Identifying disease correctly is a challenging task and requires expertise. Frequently the illnesses or its signs like colored spots or streaks can be observed on the plant leaves. The plants diseases are usually caused by microbes including fungi, bacteria, and viruses. There is a wide spectrum of signs and symptoms which differ because of the cause or etiology of the plant disease. Neural networks have been anemerging application in numerous and diverse areas as examples of end-to-end learning. The nodes in a neural network are mathematical functions that take numerical inputs from the incoming edges and provide a numerical output as an outgoing edge. The CNN may hold its applications in the agricultural field including the identification of diseases and also to quantify the diseased area. Usually, the diseases are identified by naked-eye observation by an expert. This method involves a huge time on vast farms or land. The use of convolutional neural networks in the recognition and detection of plant diseases early will be effective to increase the quality of products. To develop such a precise image classifier aimed at the diagnosis of diseases of plants, we need a large, processed and verified dataset containing various diseased and healthy plant images. Plant Village project has collected thousands of plant images and made it open and free to use. The dataset is already processed and is available in three versions as Colored, grayscale and segmented

2. LITERATURE SURVEY

2.1 Background History:

A lot of work has been devoted to detecting leaf diseases using image processing in history and it continues to attract researchers to carry out their research work in this field. Automatic crop disease detection using image processing and machine learning has been gaining in recent years.

Plant disease identification has been an important area of research and development for many years. The history of plant disease identification can be traced back to the early 20th century when scientists began studying the symptoms of plant diseases in order to understand their causes and develop methods for controlling them.

In the past, plant disease identification relied heavily on the expertise of human experts who had years of experience in recognizing the symptoms of various plant diseases. However, this process was time-consuming, expensive, and often relied on subjective interpretation of symptoms, which could lead to errors in diagnosis.

2.2 Related Work

1) "Deep learning-based crop disease recognition using convolutional neural networks" by Mohanty et al. (2016) - This paper presents a CNN-based approach for identifying crop diseases using images of leaves. The authors collected a dataset of 54,306 images of 14 Plant Disease Identifier PRPCEM, AMRAVATI\\BE (CSE)\\2022-23 5 crop species with 26 different diseases. They achieved an accuracy of 98.06% on the test set.
2) "Agricultural disease recognition using deep learning" by Kamilaris et al. (crop diseases using images. The authors used a dataset of 54,306 images of 14 crop species with 26 different diseases. They achieved an accuracy of 98.06% on the test set.
3) "Deep learning-based classification of tomato diseases using convolutional neural networks" by Sladojevic et al. (2016) - This paper proposes a CNN-based approach for classifying tomato diseases using images of leaves. The authors collected a dataset of 5,000 images of 6 different tomato diseases. They achieved anaccuracy of 99.13% on the test set.
4) "Deep learning for plant disease detection and diagnosis" by Ferentinos (2018) - This paper presents a CNN-based approach for detecting and diagnosing plant diseases using images. The author collected a dataset of 54,306 images of 14 crop species with 26 different diseases. They achieved an accuracy of 99.35%on the test set.
5) "Plant disease detection using a convolutional neural network with transfer learning" by Gangrade et al. (2019) - This paper proposes a CNN-based approach for detecting plant diseases using transfer learning. The authors used the Inception V3 network as a base model and fine-tuned it on a dataset of 54,306 images of 14 crop species with 26 different diseases. Theyachieved an accuracy of 97.7% on the test set.
6) According to paper ―Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features" [3] by S. Arivazhagan, the disease identification process includes four main steps as follows: first, a color transformation structure is taken for the input RGB image, and then by means of a specific threshold value, the green pixels aredetected and uninvolved, which is followed by segmentation process, and for obtaining beneficial segments the texture statistics are computed. Plant Disease Identifier PRPCEM, AMRAVATI\\BE (CSE)\\2022-23 6
7) Applying image processing techniques to detect plant diseases", a methodology for early and accurate plant diseases detection, using artificial neural network (ANN) and diverse image processing techniques. Kulkarni et al. in the paper ― As the proposed approach is based on ANN classifier for classification and Gabor filter for feature extraction, it gives better results witha recognition rate of up to 91%. At last, classifier is used for the features that are extracted to classify thedisease.

2.3. Limitations of Existing System:

In the paper "Deep Learning for Image-Based Plant Detection" the authors Prasanna Mohantyet al., has proposed an approach to detect disease in plants by training a convolutional neural network. The CNN model is trained to identify healthy and diseased plants of 14 species. The model achieved an accuracy of 99.35% on test set data. When using the model on images procured from trusted online sources, the model achieves an accuracy of only 31.4%, while this is better than a simple model of random selection, a more diverse set of training data can aid to increase the accuracy. Also, some other model or neural network training variations may yield higher accuracy, thus paving the path for making plant disease detection easily available to everyone.

3. PROPOSED WORK

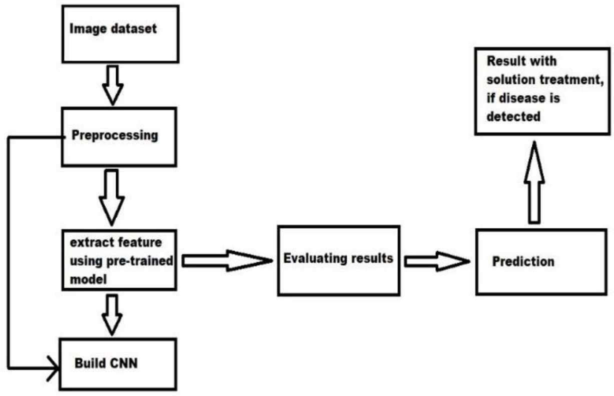

FIGURE 3.1: PROPOSED SYSTEM DESIGN.

A system design of plant disease identifier using CNN (Convolutional Neural Network) would involve the following steps:
The first step is to collect a dataset of images of healthy and diseased plants. This dataset should contain a sufficient number of images to train the CNN model. The images should be of high quality, and the diseases should be accurately labeled.

The dataset should be pre-processed by resizing the images to a fixed size and normalizing the pixel values. This will help the model to learn features more efficiently.

The next step is to train the CNN model using the pre-processed dataset. The CNN model consists of multiple layers, including convolutional, pooling, and fully connected layers.

The convolutional layers are responsible for learning features from the images, while the fully connected layers are used for classification. The model is trained by minimizing the loss function using a training algorithm such as backpropagation.

After training the model, it should be tested using a separate dataset that the model has not seen before. This will help to evaluate the accuracy of the model and identify any issues that need to be addressed.

Once the model is trained and tested, it can be deployed to identify plant diseases in real time. The system should take an image of a plant as input and output the predicted disease.

On the basis of the output given by the model, if the plant is diseased, it shows the result as a diseased plant and will also provide the solution to cure the disease i.e., the pesticide needed to get rid of the pests because of which the plant was affected. And will provide us with the products that will be needed to treat the diseased plant.

4. SYSTEM IMPLEMENTATION

4.1.1 Software Requirements:

i. TensorFlow:

TensorFlow is an open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library used for machine learning applications such as neural networks. TensorFlow is available on various platforms such as64- bit Linux, macOS, Windows, and mobile computing platforms including Android and iOS. TensorFlowcomputations are expressed as stateful dataflow graphs. The name is TensorFlow derives from operations that such neural networks perform on multidimensional data arrays, which are referred to as tensors.

pip install tensorflow – command.

ii. NumPy:

NumPy is a library of Python programming language, that adds support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate over these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several developers. In 2005 Travis Oliphant created NumPy by incorporating features of computing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors.

pip install numPy - command

iii. Keras:

Keras is an open-source neural-network library written in Python. It is capable of running on top of TensorFlow, Microsoft Cognitive Toolkit, Theano, or PlaidML. It is designed to enable fast experimentation with deep neural networks and focuses on being user-friendly, modular, and extensible.

pip install keras – command

iv. Flask:

Flask is a web framework. This means Flask provides you with tools, libraries, and technologies that allow you to build a web application. This web application can be some web pages, a blog, a wiki, or go as big as a web-based calendar application or a commercial website. Flask is part of the categories of the micro-framework. Micro-framework is normally a framework with little to no dependencies on external libraries.

pip install flask – command

v. Python:

Python: Python is a popular programming language used for machine learning and deep learning projects. We used Python to implement the CNN model for plant disease identification in this project.

vi. Jupyter Notebook:

Jupyter Notebook is an open-source web application for creating and sharing documents containing live code, equations, visualizations, and narrative text. We used Jupyter Notebook in this project to write and run the Python code.

These software requirements are all freely available and can be easily installed using package managers like pip or conda. Additionally, some cloud-based platforms like Google Colab provide pre-installed environments with these libraries for machine learning and deep learning projects.

4.1.2 Hardware Requirements:

i. Memory:

Python uses a portion of the memory for internal use and non-object memory. Another part of the memory is used for Python objects such as int, dict, list, etc. Python contains the object allocator that allocates memory within the object area. The object allocator gets a call every time the new object needs space. Memory management requires that the programmer provides ways to dynamically allocate portions of memory to programs, when requested, and free it for
reuse when it is no longer needed. This is critical in any advanced computer system, where more than a single process might be running at any given time.

ii. Processor:

The processor, also known as the CPU, provides the instructionsand processing power the computer needs to do its work. The more powerful and updated your processor, the faster your computer can complete its tasks. The CPU is seen as the main and most crucial integrated circuitry (IC) chip in a computer, as it interprets most computer commands

4.2 Implementation Details:

A. Installation:

1) Install Jupyter Notebook (Google Colab) for model training. Used Spyder for model deployment on the local system. To use Jupyter Notebook and Spyder, just install anaconda.

2) Download and install Python version 3 from the official Python Language websitehttps://python.org

3) Install the following dependencies via pip:

- TensorFlow
- Keras

- NumPy
- Flask

B. Implementation:

The implementation details of a plant disease identifier project using CNNs can be divided into several steps, which are:

1. The very first step is to Import all the necessary libraries.

```
In [9]: #import Libraries
        import keras
        from keras.preprocessing.image import ImageDataGenerator
        from keras.optimizers import Adam
        from keras.callbacks import ModelCheckpoint
        import matplotlib.pyplot as plt
```

Figure 4.1: Imported Libraries.

2. To Train the Dataset, set the data file path and then imported the file in code.
   The dataset used in the project was referred from a programmer name "Akash Jade".
3. Pre-processing and Augmentation of the collected dataset are done using pre-processing.

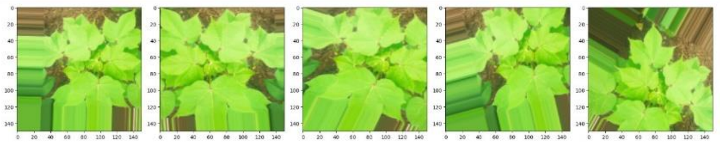

Figure 4.2: Pre-processing of Images.

4. Building CNN (Convolutional Neural Network) Model for the classification of plant diseases.
5. Then Compile and Train the CNN model.

How CNN Works:

A CNN is composed of stacking of several building blocks: convolution layers, pooling layers (e.g., max pooling), and fully connected (FC) layers. A model's performance under particular kernels and weights is calculated with a loss function through forward propagation on a training dataset, and learnable parameters, i.e., kernels and weights, are updated according to the loss value through backpropagation with a gradient descent optimization algorithm. ReLU rectified linear unit.

CNN can have multiple layers, each of which *learns* to detect the different features of an input image. A filter or kernel is applied to each image to produce an output that gets progressively better and more detailed after each layer. In the lower layers, the filters can start as simple features.

At each successive layer, the filters increase in complexity to check and identify features that uniquely represent the input object. Thus, the output of each convolved image -- the partially recognized image after each layer -- becomes the input for the next layer. In the last layer, which is an FC layer, the CNN recognizes the image or the object it represents.

With convolution, the input image goes through a set of these filters. As each filter activates certain features from the image, it does its work and passes on its output to the filter in the next layer. Each layer learns to identify different features and the operations end up being repeated for dozens, hundreds or even thousands of layers. Finally, all the image data progressing through the CNN's multiple layers allow the CNN to identify the entire object.

```
61/61 [==============================] - 86s 1s/step - loss: 0.3291 - accuracy: 0.8847 - val_loss:
0.3448 - val_accuracy: 0.8364
Epoch 44/500
61/61 [==============================] - ETA: 0s - loss: 0.3307 - accuracy: 0.8713
Epoch 44: val_accuracy did not improve from 0.91049
61/61 [==============================] - 88s 1s/step - loss: 0.3307 - accuracy: 0.8713 - val_loss:
0.3088 - val_accuracy: 0.8673
Epoch 45/500
61/61 [==============================] - ETA: 0s - loss: 0.3339 - accuracy: 0.8698
Epoch 45: val_accuracy did not improve from 0.91049
61/61 [==============================] - 88s 1s/step - loss: 0.3339 - accuracy: 0.8698 - val_loss:
0.3195 - val_accuracy: 0.8673
Epoch 46/500
61/61 [==============================] - ETA: 0s - loss: 0.3062 - accuracy: 0.8903
Epoch 46: val_accuracy did not improve from 0.91049
61/61 [==============================] - 82s 1s/step - loss: 0.3062 - accuracy: 0.8903 - val_loss:
0.2604 - val_accuracy: 0.8981
Epoch 47/500
 7/61 [==>...........................] - ETA: 1:14 - loss: 0.3385 - accuracy: 0.8839
```

Figure 4.3: Validation Accuracy and Loss of the Model.

6. Open Spyder and create a new project then create folders and files according to below hierarchy of the project.

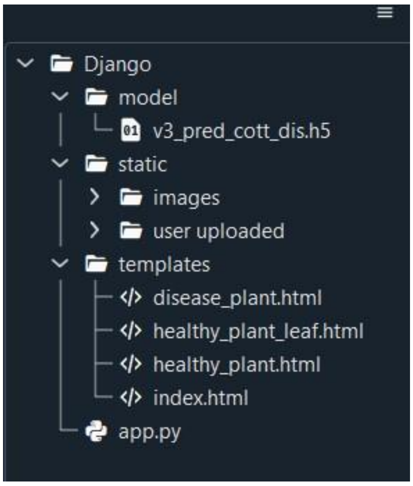

Figure 4.4: Hierarchy of Folders.

4.3. System Execution Details:

Run the "app.py" file in Spyder or you can run it using the Anaconda prompt. Then you will get a localhost address like "http://127.0.0.1:5000/ "Enter it in any browser in your system it will lead us to the front webpage of the project.

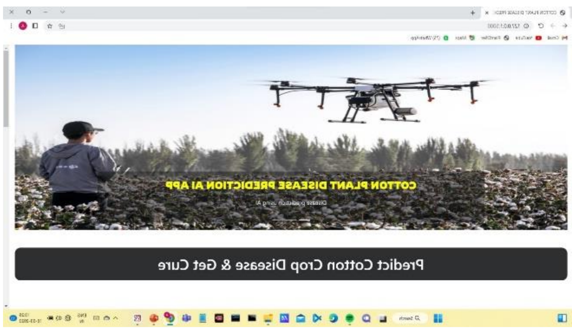

Figure 4.5: Webpage.

After opening the website's first page, we will find the "choose file" button from which we can choose the picture of a diseased plant.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

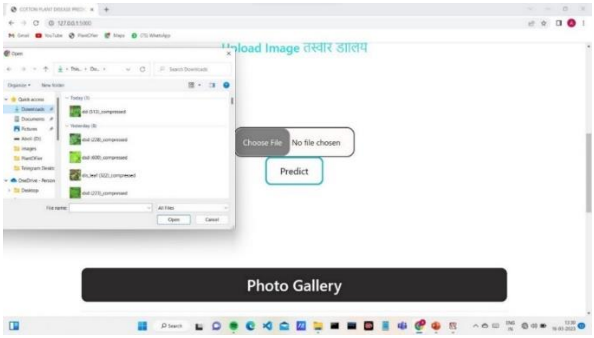

Figure 4.6: Selecting a Picture of a Diseased Plant.

After selecting the desired picture of the plant press the predict button and it will show the prediction if the selected image of the plant is diseased or healthy.

Figure 4.7: Prediction of Diseased Plant.

After predicting the disease of the plant, it will provide us with the solution needed to cure the plant and a suitable product will also be suggested.

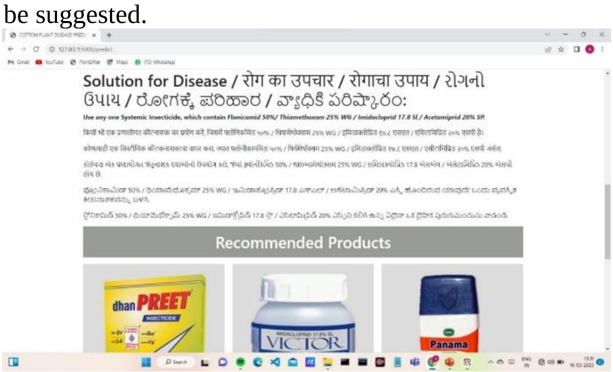

Figure 4.8: Solution and Product Recommendation. For Diseased Plant.

5. RESULT ANALYSIS

As expected, our algorithm detects the disease of the plant with utmost accuracy and also provides the solution needed to cure the disease of the plant with product recommendations.

Compared to the previous projects related to Plant Disease Identification, recent plant disease identifier projects using CNNs have several advantages over earlier projects that did not utilize deep learning techniques:

Improved accuracy: CNNs are highly effective in identifying patterns and features in images, which makes them well-suited for identifying plant diseases. As a result, recent projects that use CNNs tend to have higher accuracy rates compared to earlier projects that relied on traditional machine learning or computer vision techniques.

Speed and efficiency: CNNs are designed to process large amounts of data quickly and efficiently, which makes them idealfor use in plant disease identification. Recent projects using CNNs have been shown to be significantly faster than earlier projects, which can help save time and resources.

Transfer learning: Recent projects using CNNs often utilize transfer learning, which involves using pre-trained models as a starting point and fine-tuning them for the specific task of plant disease identification. This approach can significantly reduce the amount of time and resources needed to train the model from scratch.

Overall, recent plant disease identifier projects using CNNs represent a significant improvement over earlier projects in terms of accuracy, speed, and accessibility. They have the potential to greatly improve crop yields and reduce the impact of plant diseases on agriculture.

6. ADVANTAGES & LIMITATAIONS

The plant disease identifier project using CNN has several advantages, including:

i. Increased accuracy: The use of deep learning models, specifically CNNs, has shown promisingresults in plant disease identification. The CNN architecture is capable of learning complex features from images, which can lead to improved accuracy in disease identification.

ii. Reduced human error: Traditional methods of disease identification often rely on human expertise, which can lead to errors due to misinterpretation or misdiagnosis. The use of an automated system can reduce the likelihood of human error, leading to more reliable and accurate disease identification.

iii. Faster diagnosis: Automated disease identification systems can provide a faster diagnosis of plant diseases compared to traditional methods, which can take days or weeks to identify the disease. A faster diagnosis can lead to timely treatment, reducing the spread of the disease and preventing crop losses.

iv. Cost-effective: The use of an automated disease identification system can be cost-effective compared to traditional methods, which may require specialized equipment or expert consultation. The plant disease identifier project using CNN can be deployed as a web application, making it easily accessible and cost-effective for farmers and other stakeholders in the agriculture industry.

v. Scalability: The automated system can be scaled up to handle large datasets and can be deployed in various locations, making it a valuable tool for plant disease identification across different regions and crops.

Limitations:

i. The accuracy of the CNN model heavily depends on the quality and quantity of the training data. If the training dataset is limited or biased, it may negatively affect the performance of the model, leading to misdiagnosis.

ii. The CNN model is a black box, meaning it is difficult to understand how the model is making its predictions. This lack of interpretability can make it challenging to identify the reasons for incorrect diagnoses or to adjust the model'sbehavior.

7. CONCLUSION

This paper proposes a CNN-based method for plant disease classification using the leaves of diseased plants. Building such a neural network with high efficiency is a complex task. Transfer learning can be employed to achieve greater efficiency. Inception v3 is one of the models available that inherently can classify images and can be trained to identify different classes. Thus, Inception v3 can play a key role in obtaining fast and effective plant disease identifiers. Also, by dataset classification using the contour method, the training set can be chosen to ensure proper training of the model for all features. This provides better feature extraction than randomly classifying the dataset. Optimal results were obtained by employing the methods specified in the paper. Thus, with the implementation and use of these methods for plant disease classification losses in agriculture can be reduced.

The plant disease identifier project using CNN is a promising tool for the agriculture industry, helping to improve the accuracy and speed of plant disease identification. The use of deep learning models, specifically CNNs, has shown promising results in identifying plant diseases from images. This project has several advantages, including increased accuracy, reduced human error, faster diagnosis, cost-effectiveness, and scalability

REFERENCES

[1] Mohanty, S. P., Hughes, D. P., & Salathia, M. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7(September), [1419] *https://doi.org/10.3389/fpls.2016.01419* (2016).

[2] Santhosh Kumar S, and B. K. Raghavendra. Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review. 2019 5th *International Conference on Advanced Computing & Communication Systems (ICACCS)* (2019).

[3] Jyotsna Bankar, and Nitin R Gavai. Convolutional Neural Network based Inception v3 Model for Animal Classification. *International Journal of AdvancedResearch in Computer and Communication Engineering*, (May 2018).

[4] Indira. Abirami Devaraj, Karunya Rathan, Sarvepalli Jahnavi, and K Identification of Plant Disease using Image Processing Technique. 2019 *International Conference on Communication and Signal Processing (ICCSP)* (2019).

[5] Ser Serawork Wallelign, Mihai Polceanu, Cédric Buche. Soybean Plant Disease Identification Using Convolutional Neural Network. *FLAIRS-31*, May 2018, Melbourne, United States. *https://hal.archives-ouvertes.fr/hal-01807760*(May 2018).

[6] S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini. 2013.Detection of the unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agric Eng. Int: CIGR Journal* (2013).

[7] Islam, M. T., Xie, X., & Yang, C. Leaf disease detection and classification using deep learning convolutional neural network. Symmetry, 11(8), 1016.(2019).

[8] Singh, D., Singh, R., & Kaur, M. Plant disease identification using convolution neural networks. *International Journal of Advanced Science and Technology*, 29(7), 1476-1484. (2020)

[9] Chakraborty, D., Sarkar, S., & Nasipuri, M. A *systematic review on plant disease detection using deep learning techniques*. Plant Disease, 104(3), 622-641. (2020)

[10] Kavitha, C., & Arockiam, L. (2019). Detection of plant diseases using convolutional neural network. *International Journal of Recent Technology and Engineering*, 7(S6), 100-103.