**PLANT DISEASE DETECTION USING DEEP LEARNING**

**ABSTRACT:**

Conventional techniques of plant health monitoring have absence of formulating on time complex decisions subjective to intelligence and skill set in comprehending plant pathology (Aydin Kaya et al, 2019). Keeping an eye on plant health in a planned manner is decisive to foresee the agricultural yield and post-harvest losses in the industry. It goes without saying that, early bird catches the worm. Hence, timely detection and diagnosis of plant diseases is paramount to chain the disease spread. Furthermore, it maintains sustainability in agro-economic sector (Konstantinos P. Ferentinos et al, 2018) as large fraction of country’s population depends on agriculture for its livelihood. With the advent of technologies for instance remote sensing, computer vision and machine learning are offering reliable, accurate and precise solutions to assess plant quality in real time (Edna Chebet Too et al, 2018). Precision agriculture or satellite farming aids farmers to make intelligent decisions and earn optimal returns on invested inputs while reserving resources.

**INTRODUCTION:**

A plant disease is defined as abnormal appearance of a plant which reduces its productivity or utility to mankind. The major economic havoc in agricultural sector is contributed by crop damage happening due to pathogens spread across varied seasons (Mohanty et al, 2018). Virus, fungi, bacterium and protozoa are infectious microorganisms downgrading the crop quality extensively (Fig 1) (Artzai Picon et al, 2019). Nevertheless, Global warming along with recurring climate changes are equally responsible for inferior quality and shrink quantity of crop production.

The event of plant infections effect sly affects agricultural creation, and if the plant infections are not identified in time, there will be an expansion in food uncertainty . Specifically, the primary yields, for example, rice, maize, and so forth., are basic for ensuring the food gracefully and agricultural creation. The early admonition and conjecture are the premise of powerful anticipation and control for plant maladies. They assume vital jobs in the administration and dynamic for farming creation. As of not long ago, be that as it may, the visual perceptions of experienced makers are as yet the essential methodology for plant disease discovery in country regions of creating nations; this requires nonstop observing of specialists, which may be restrictively costly in enormous homesteads . In addition, in some far off zones, farmers may need to go significant distances to contact specialists, which makes the counseling excessively costly and tedious. In any case, this methodology must be done in constrained zones also, can't be very much broadened. Programmed acknowledgment of plant maladies is a fundamental examination theme, as it might demonstrate benefits in observing huge fields of harvests, and subsequently consequently recognize the side effects of diseasees when they show up on plant leaves . Thusly, searching for a quick, programmed, more affordable, and exact technique to perform plant malady recognition is of extraordinary reasonable criticalness.

Recent advancements in the field of image processing have blaze a trail for the researchers and scientists to contribute significantly to agricultural revolution. Moreover, fast embedded processors such as Graphical Processing Units (GPU) are offering viable and fast development platforms by virtue of machine learning techniques on multiple platforms (Haseeb Nazki et al, 2020). As a result, open source software packages are available in discrete languages for cross platform implementation of machine learning algorithms in a subtle manner (Saptarshi Sengupta et al, 2020). Deep learning technique has crop up as clear winner in its category. It has been exhaustively probed in diverse application spheres which incorporates agriculture, health, automobile and sentiment analysis (Kamal KC et al, 2019). Deep Convolution Neural Network (CNN) has taken a leap forward in the field image processing and segmentation. Distinct CNN architectures including VGG Net, GoogLeNet, ZFNet, AlexNet, LeNet-5, Inception V4, ResNet and DenseNet came to life to deal with the problems related to training process degradation (Edna Chebet Too et al, 2018, Saptarshi Sengupta et al, 2020). The training process degradation arises as a result of adding more layers to network design in a ray of hope of getting refined accuracy. The deep convolution neural networks (CNNs) can be vital in such situations due to their capability of consuming different type of inputs at various stages of training process leading to optimized performance. CNN is a flexible framework capable of performing feature selection and classification tasks in close integration yielding maximum throughput (Artzai Picon et al 2019).

**Research Questions:**

The main research questions that will be addressed in this work is as follows:

* How to use computer vision and machine learning for plant disease classification?
* How to develop an accurate and efficient method for diagnosis of plant diseases?

**BACKGROUND:**

Plants get various diseases which diminishes their efficiency. These diseases can be recognized by their shading, surface, and shape. Maladies of plants can happen in leaves stem. These ailments can cause because of infection, organisms, microbes. They can be communicated through insects. Farmers are qualified pundits of the ailments, as a result of their encounters. Be that as it may, truth be told they are not exact and right some of the time. On the off chance that farmers cannot control the disease, at that point they call the specialists, yet this can be tedious. Generally, the infection can be on leaves and on stem of plants. These diseases can be viral, contagious, bacterial diseases tainted by creepy crawlies, nematodes, rust on plants. It is significant obligation for farmers to gain proficiency with these diseases early. The image based characterization of plant diseases is a troublesome supportive of problem with a wide assortment of difficulties related, including the presence of manifestations with broad scope of visual attributes, possibility of different concurrent issues in a solitary plant, and various issues having comparative side effects, among others. The image based arrangement of plant infections is a troublesome supportive of problem with a wide assortment of difficulties related, including the presence of indications with broad scope of visual qualities, possibility of different synchronous issues in a solitary plant, and various issues having comparative manifestations, among others. Extraneous factors, for example, impedance brought about by the image foundation and light varieties related to catch conditions add considerably greater unpredictability to the issue. While the blend of image handling and AI has prompted numerous advances, useful utilization of apparatuses like these has been constrained. In the most recent couple of years, a few investigations have utilized the ideas of profound learning, and Convolutional Neural Networks (CNN) specifically, to attempt and make this sort of hardware more precise.

I have started working on the project and have mainly done the study part. I have studied about the various of plant diseases that are displayed in plant leaf images. These diseases can be categorized as bacterial diseases like aster yellows, bacterial wilt, scab etc. Types of common fungal disease: chestnut blight, late blight, anthracnose etc. Types of common virus disease, curly top, mosaic etc. The following figure shows the leaf images with diseases.

   

One of the principle challenges in AI identified with the agricultural imaging space is the manner by which to adapt to the little datasets and predetermined number of commented on tests, explicitly while utilizing regulated AI calculations that need named information and enormous number of preparing models. Gathering plant infection related information is an intricate and costly methodology and requires the coordinated effort of individuals from various fields at differentiating stages. Albeit open datasets are accessible, most datasets are as yet restricted in size and material to explicit issues. I have also searched about the datasets that are available for carrying out the classification process. Some of the commonly used datasets are explored for collection of the data. They are Leaf, Folio Leaf from UCI and the rest three are PlantVillage, Flavia and Swedish datasets. The Leaf dataset constitutes of 40 different types of plant species which further incorporates several leaf specimens corresponding to each species. Some of them are simple leaves and the rest are complex leaves. The PlantVillage dataset contains 54306 leaf images spread across 38 different categories of plant species. It is one of most popularly used publically available dataset for testing the performance of various machine learning algorithms. The next dataset considered for experimentation is Folio Leaf dataset which contains dual mixture of simple and complex leave images taken in bright sunlight. The last two datasets namely Swedish and Flavia contains numerous snapshots of plant images taken from different angles.

The dataset that will be used in this work will be the plant village dataset.

|  |  |  |
| --- | --- | --- |
| 1. Pepper\_\_bell\_\_\_Bacterial\_spot | 1. Potato\_\_\_Early\_blight | 1. Tomato\_Bacterial\_spot |
| 1. Tomato\_Leaf\_Mold | 1. Pepper\_\_bell\_\_\_healthy | 1. Potato\_\_\_healthy |

The dataset is identified and the study of the python modules to be used is completed. The work on python part is also started. I have started with the exploratory data analysis of the dataset to get a better understanding. In this the python codes are written to read the data and view the data and the various statistics. This information will be useful when the model will be developed and evaluated. The image processing methods are also applied to see what kind of results are obtained. Like the canny edge detector is applied to find the edges of the images. Canny edge detection technique is used for finding the venation example of the leaves. In the main phase of this strategy the leaf images are smoothed so as to expel the commotion. Next stage includes finding the slopes with high greatness by nearby maxima for recognizing veins. At that point veins are featured where high estimations of spatial subsidiaries are found. Edge up-and-comers are chosen after twofold thresholding. At long last veins are recognized by killing veins that are not associated with solid veins.

**METHODS:**

The deep convolution neural networks (CNNs) can be indispensable in such circumstances because of their ability of devouring diverse sort of contributions at different phases of preparing process prompting advanced execution. CNN is an adaptable system equipped for performing feature choice and order undertakings in close coordination yielding most extreme throughput. The objective of the project is to design a deep convolution network for classification of leaves based on plant disease.

Research in agriculture space is pointed towards increment the quality and amount of the item at less consumption with more benefit. The nature of the rural item might be corrupted because of plant ailments. These infections are brought about by microbes viz.., growths, microorganisms and infections. In this way, to distinguish and group the plant disease in beginning phase is a huge assignment. Farmers require steady observing of specialists which may be restrictively costly and tedious. Initially, the examples of plant disease images are gathered and marked dependent on the information of specialists in the field. At that point, the image preparing strategies including dark change, image sifting, image honing and resizing, and so forth., are performed on the procured images, and new example images are produced to advance the dataset utilizing the information growth techniques. For instance, irregular turn, flipping, and interpretation are used to grow the dataset. From that point forward, the example images are contribution to the proposed technique for model preparing. Consequently, the prepared model is applied for the class forecast of inconspicuous images, and the consequences of plant malady recognizable proof are acquired in the long run. The point by point depictions of these stages are represented in resulting areas.

Contingent upon the applications, numerous frameworks have been proposed to tackle or if nothing else to diminish the issues, by making use of image processing and machine learning. The proposed work includes four modules:

* Leaf dataset collection
* Feature extraction,
* Classification of diseased leaves.
* Evaluation of the developed method

Prior to this work image pre-processing is also required. Image preprocessing is to expel commotion from the image or other item expulsion, diverse preprocessing procedures. Here we are utilizing image scaling. Image scaling is utilized to convert the first image into thumbnails in light of the fact that the pixel size of unique image is huge and it require more opportunity for the general procedure consequently subsequent to changing over the image into thumbnails the pixel size will get diminishes and it will require less time. Image division is one of the for the most part utilized strategies to arrange the pixels of a image accurately in a decision oriented application. It partitions a image into various discrete districts to such an extent that the pixels have high likeness in every district and high differentiation between areas. Feature Extraction is a significant part in the malady identification. It assumes a significant job in ID of an object. Feature extraction is utilized in numerous applications in image preparing. Shading, surface edges, morphology are the features, which are utilized in ailment detection. The last stage is the location of the maladies and with the help of ailment characterize the plants with the malady matches with the given dataset. For the disease location and characterization, we are executing the profound learning calculation. Profound learning calculation is utilized to characterize the indicated image into fitting disease thus it will be simple to identify the disease and discover the cure over the disease.

All the above steps will be carried out using Python programming language. The keras module in python will be used for implementation of the classification. Keras is a model-level library, giving elevated level structure squares to creating profound learning models. It doesn't deal with low-level activities, for example, tensor control and separation. Rather, it depends on a particular, all around advanced tensor library to do as such, filling in as the backend motor of Keras. As opposed to picking a solitary tensor library and binds the execution of Keras to that library, Keras handles the issue in a measured manner therefore a few diverse backend motors can be stopped consistently into Keras. As of now, the three existing backend usage are the TensorFlow backend, the Theano backend, and the Microsoft Cognitive Toolbox (CNTK) backend. Later on, almost certainly, Keras will be stretched out to work with significantly more profound learning execution motors. There are two different ways to characterize a model: utilizing the Sequential class (just for direct heaps of layers, which is the most widely recognized system design by a long shot) or the utilitarian API (for coordinated non-cyclic diagrams of layers, which lets you fabricate totally self-assertive structures).

The jupyter notebook will be used for the keras module. A notebook is a record created by the Jupyter Notebook application (https://jupyter.org), which you can alter in your program. It blends the capacity to execute Python code with rich content editing capacities for commenting on what you're doing. A scratch pad additionally permits you to separate long tests into little pieces that can be executed autonomously, which makes advancement intelligent and implies you don't need to rerun the entirety of your past code if something turns out badly late in an analyses. Since the work will require GPU processing power thus cloud will be used for implementation. Therefore, for using GPU power Google colab notebook will be used for implementation.

**REFERENCES:**

1. Sengupta, S., Basak, S., Saikia, P., Paul, S., Tsalavoutis, V., Atiah, F., Ravi, V. and Peters, A., 2020. A review of deep learning with special emphasis on architectures, applications and recent trends. *Knowledge-Based Systems*, p.105596.
2. Kamal, K.C., Yin, Z., Wu, M. and Wu, Z., 2019. Depthwise separable convolution architectures for plant disease classification. *Computers and Electronics in Agriculture*, *165*, p.104948.
3. Amara, J., Bouaziz, B. and Algergawy, A., 2017. A deep learning-based approach for banana leaf diseases classification. *Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband*.
4. Brahimi, M., Boukhalfa, K. and Moussaoui, A., 2017. Deep learning for tomato diseases: classification and symptoms visualization. *Applied Artificial Intelligence*, *31*(4), pp.299-315.
5. Lee, S.H., Chan, C.S., Mayo, S.J. and Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition*, *71*, pp.1-13.
6. Murat, M., Chang, S.W., Abu, A., Yap, H.J. and Yong, K.T., 2017. Automated classification of tropical shrub species: a hybrid of leaf shape and machine learning approach. *PeerJ*, *5*, p.e3792.
7. DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E.L., Yosinski, J., Gore, M.A., Nelson, R.J. and Lipson, H., 2017. Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology*, *107*(11), pp.1426-1432.
8. Cruz, A.C., Luvisi, A., De Bellis, L. and Ampatzidis, Y., 2017. X-FIDO: An effective application for detecting olive quick decline syndrome with deep learning and data fusion. *Frontiers in plant science*, *8*, p.1741.
9. Liu, B., Zhang, Y., He, D. and Li, Y., 2018. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, *10*(1), p.11.
10. Oppenheim, D. and Shani, G., 2017. Potato disease classification using convolution neural networks. *Advances in Animal Biosciences*, *8*(2), pp.244-249.
11. Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and electronics in agriculture*, *153*, pp.46-53.
12. Dong, X., Chen, S. and Pan, S., 2017. Learning to prune deep neural networks via layer-wise optimal brain surgeon. In *Advances in Neural Information Processing Systems* (pp. 4857-4867).
13. Han, S., Mao, H. and Dally, W.J., 2015. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *arXiv preprint arXiv:1510.00149*.