Disease Detection in Plant using Machine Learning

Sudarshan Barure[1], Bhushan Mahadik[2], Monali Thorat[3], Akhil Kalal[4]

[1,2,3,4]Department of Information Technology, *Jayawantrao Sawant College of Engineering, Savitribai Phule Pune University, Maharashtra, India*

Abstract - *The paper presents plant disease diagnosis using image processing techniques for automated vision system used in agricultural field. In agriculture research of automatic leaf disease detection is essential one in monitoring large fields of crops, and thus automatically detects symptoms of disease as soon as they appear on plant leaves. The proposed decision making system utilizes image content characterization and supervised classifier type back propagation with feed forward neural network. Image processing techniques for this kind of decision analysis involves preprocessing, feature extraction and classification stage. The system will be used to classify the test images automatically to decide plant either abnormality or good one. For this approach, automatic classifier CNN will be used for classification based on learning with some training samples of that two category*.

***Key Words*: Computational Neural Network, Machine Learning, Disease Detection, Image processing.**

1. INTRODUCTION

The agriculturist in provincial regions may think that it's hard to differentiate the malady which may be available in their harvests. It's not moderate for them to go to agribusiness office and discover what the infection may be. Our principle objective is to distinguish the illness introduce in a plant by watching its morphology by picture handling and machine Learning. Pests and Diseases results in the destruction of crops or part of the plant resulting in decreased food production leading to food insecurity. Also, knowledge about the pest management or control and diseases are less in various less developed countries. Toxic pathogens, poor disease control, drastic climate changes are one of the key factors which arises in dwindled food production.

Various modern technologies have emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize the productivity. Various Laboratory based approaches such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. However, these techniques are not cost effective and are high time consuming. In recent times, server based and mobile based approach for disease identification has been employed for disease identification. Several factors of these technologies being high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition.

2. RELATED WORK

Sr.No	Journal Name And Publication Year	Title	Technolog y	Identify Diseases
1	International Conference On internet of Things and Intelligence System Analysis[IEEE][2018]	KrishiMitr (Farmer Friend): Using Machine Learning to Identify Diseases in Plants.	Tensor Flow framework by CCN model	Calculate CNN Model.
2	International Conference for Convergence in Technology[IEEE] [2018]	Mango leaf Deficiency Detection using digital image processing and Machine learning	Image Processing	Calculate only leaf area.
3	International Research Journal of Engineering and Technology[IEEE] [2018]	Plant Disease Detection Using Machine Learning	Canny Edge Detection Algorithm.	Diseases in crop. Mostly on leaves.

Chart -1: Comparison Table of Literature **Survey**

3. METHODOLOGY

Image Dataset

One source of images for the dataset is the Plant Village database, prepared by Hughes et al.[6], which contains more than 50,000 images of more than 35 diseases of 16 plant types. Another field-based dataset of 7,000 images was collected from fields across north-west India for 4 plant types. Lastly, a total of 15,000 images were collected from Internet. Special care was taken to ensure random backgrounds and different stages of infection / damage.

4. SYSTEM ARCHITECTURE

The paper presents plant disease detection using image processing techniques for automated vision system used at agricultural Field. In agriculture research of automatic leaf disease detection is essential one in monitoring large fields of crops, and thus automatically detects symptoms of disease as soon as they appear on plant leaves. For this approach, automatic classifier CNN will be used for classification based on learning with some training samples of that two category. Finally, the simulated result shows that used network classifier provides minimum error during training and better accuracy in classification. Farmer approach to agricultural consultant for recognition of plant Disease. At Processing, an input image will be resized and region of interest selection performed if needed. Here, color and texture features are extracted from an input for network training and classification. Color features like mean, standard deviation of HSV color space and texture features like energy, contrast, homogeneity and correlation. The system will be used to classify the test images automatically to decide plant either abnormality or good one. For this approach, automatic classifier CNN will be used for classification based on learning with some training samples of that two category. Finally, the simulated result shows that used network classifier provides minimum error during training and better accuracy in classification.

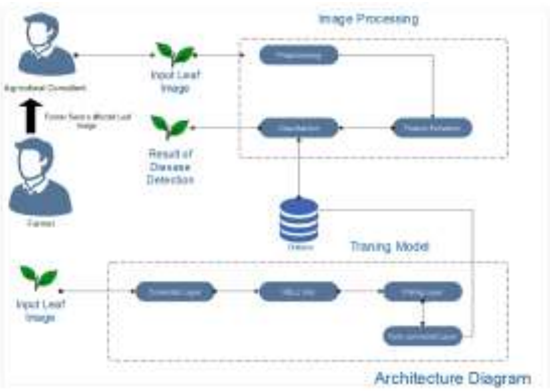

Fig.1: Architecture of System

4.1 Convolution Layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel. The convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is called \Feature Map "Padding Sometimes filter does not fit perfectly fit the input image. We have two options:

Pad the picture with zeros (zero-padding) so that it its Drop the part of the image where the filter did not fit. This is called valid padding which keeps only valid part of the image.

4.2 RELU Unit

In a neural network, the activation function is responsible for transforming the summed weighted input from the node into the activation of the node or output for that input. The rectified linear activation function is a piecewise linear function that will output the input directly if is positive, otherwise, it will output zero. It has become the default activation function for many types of neural networks because a model that uses it is easier to train and often achieves better performance.

4.3 Poling layer

Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or down sampling which reduces the dimensionality of each map but retains important information. Spatial pooling can be of different types:

Max Pooling
Average Pooling
Sum Pooling

Max pooling takes the largest element from the rectified feature map. Taking the Largest element could also take the average pooling. Sum of all elements in the Feature map call as sum pooling.

4.4 Fully Connected Layer

Fully connected layers are an essential component of Convolutional Neural Networks (CNNs), which have been proven very successful in recognizing and classifying images for computer vision. The CNN process begins with convolution and pooling, breaking down the image into features, and analyzing them independently. The result of this process feeds into a fully connected neural network structure that drives the final classification decision.

Fully connected input layer (flatten) takes the output of the previous layers, \at-tens" them and turns them into a single vector that can be an input for the next stage.

The first fully connected layer takes the inputs from the feature analysis and applies weights to predict the correct label. Fully connected output layer gives the final probabilities for each label.

4.5 Preprocessing

Image processing is divided into analogue image processing and digital image processing.

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data | the aim of digital Image processing is to improve the image data (features) by suppressing unwanted distortions and/or enhancement of some important image features so that our AI Computer Vision models can benefit from this improved data to work on.

4.6 Feature Extraction

A CNN is composed of two basic parts of feature extraction and classification. Feature extraction includes several convolution layers followed by max-pooling and an activation function. In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values (features) Intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human.

4.7 Classification

A classification is a process of placing each individual under study in many classes. Classification helps to analyses the measurements of an object to identify the Classes to which that Image belongs. To establish an efficient relation, analysts Use data.

5. Training Testing Algorithm

Input: Providing leaf image
Output: Classification of review into Healthy, Unhealthy and Neutral.

Step 1: Start
Step 2: Prepare Database (Healthy /Unhealthy)
Step 3: Preprocessing Normalization (Size of 64 X 64)
Step 4: Train CNN
Step 5: Real Image from camera/Pc
Step 6: Preprocessing (Size =64 X 64)
Step 7: Test Network
Step 8: if probability of healthy > probability of unhealthy
Display Healthy Image
Otherwise
Display Unhealthy image
Step 9: Go to Step 4
Step 10: End.

6. RESULTS AND DISCUSSION

Out of the total images, 70% were used for training the model and the rest 30% were used For validation. Figure 3 shows the evolution of accuracy and cross entropy for training and validation,

Respectively during training of the CNN model for all the plant types. Table II shows the final accuracy and cross entropy for training and validation for all the plant types. In all cases, validation accuracy greater than 94 % was achieved. Figure 4 shows an example of images of healthy and diseased leaves for wheat. The system user interface requires the user to select the plant type before detecting the disease type. For testing, author was deployed on a windows-based handheld Computer with a touchscreen interface and a sensor for image acquisition. Whereas programs Such as Assess [8] and Leaf Doctor [1] focus on quantification of severity of diseases, this focuses on detection of the type of disease so that farmers can take adequate remedial measures.

7. CONCLUSIONS

We describe a deep convolutional neural net-work based model for detection of foliar diseases in plants .With help of machine learning we emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize the productivity. Various Laboratory based approaches such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. These technologies being high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition. The given system uses resizing, threes holding and Gaussian filtering for image preprocessing. To segment the leaf area' then finally CNN classification technique is used to detect the type of plant disease.

ACKNOWLEDGEMENT

We take this opportunity to thank our paper guide Prof. A.K.Gupta and Head of the Department Dr.S.V.Todkari for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this project report. We are thankful to our Principal Dr. R.D.Kanphade and to all the staff members of the Department of Information Technology of Jayawantrao Sawant college of Engineering, Pune for their valuable time, support, comments, suggestions and persuasion.

We would also like to thank the institute for providing the required facilities, Internet access and important books.

REFERENCES

[1] S. J. Pethybridge and S. C. Nelson, "Leaf Doctor: A New Portable Application for Quantifying Plant Disease Severity," Plant Disease, vol. 99, pp. 1310–1316, 2015.

[2] J. G. A. Barbedo, L. V. Koenigkan and T. T. Santos, "Identifying multiple plant diseases using digital image processing," Biosystems Engineering, vol. 147, pp. 104–116, 2016.

[3] M. A. Clement, T. Verfaille, C. Lormel and B. Jaloux, "new colour vision system to quantify automatically foliarn discolouration caused by insect pests feeding on leaf cells," Biosystems Engineering, vol. 133, pp. 128–140,2015.

[4] https://www.overleaf.com/project/5df068b754809000 010df86f
https://pymbook.readthedocs.io/en/latest/flask.

[5] https://becominghuman.ai/image-data-pre-processing-for-neural-networks498289068258

[6] https://towardsdatascience.com/cnn-application-on-structured-dataautomated-feature-extraction-8f2cd28d9a7e