IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Cotton Leaf Disease Detection using CNN

[1]Varun Suryawanshi, [2] Yash Bhamare, [3] Rahul Badgujar, [4] Komal Chaudhary,[5] Mr. Bhushan Nandwalkar

Department of Computer Enginering

Shri Vile Parle Kelavani Mandal's institute Of Technology, Dhule, India

Abstract: India is the second largest in population and many crops Indian farmers can cultivate and Most of the farmers cultivate cotton in large numbers but the cotton leaf disease is the major problem in the past few decades and that results in a loss of crops, their productivity and money as well. The Cotton leaves are affected by the disease named Cercospora, Bacterial blight, Ascochyta blight, and Target spot. General observation by farmers may be time-consuming, expensive and sometimes inaccurate. The Cotton leaf Disease Detection and identifying the disease at an early stage is a very difficult task for the farmers. If the infection or disease on the crops was not identified by the farmers at the initial level then it will be harmful to the crops as well as for farmers. The main purpose of farming is to yield healthy crops with none disease present. It's very difficult to visually presume the health of cotton leaf. To beat this problem, a machine learning based approach is proposed which can assess the image of the leaf of the plant and detect the disease and therefore the quality of the cotton using machine learning approach. For availing this user got to upload the image then with the assistance of image processing we can get a digitized colour image of a diseased leaf then we can proceed with applying CNN to predict cotton leaf disease. Neural Network and CNN. Previously image classification algorithms like face recognition got to concentrate to where the face is found in a picture this major problem is overcome by CNN also as features of a picture are deeply processed at each layer. Every disease on a crop has different features which are extracted at each layer of the convolution network. The goal of this application is to develop a system that recognizes crop diseases. During this user has got to upload a picture on the system, Image processing starts with the digitized color image of the diseased leaf. Finally by applying the CNN disease are often predicted. The detection of plant disease may be a vital factor to stop a significant outbreak. Most plant diseases are caused by fungi, bacteria, and viruses. Traditionally farmer visually checks the disease. This paper presents an approach for careful detection of diseases and timely handling to stop the crops from heavy losses. The diseases on cotton are a critical issue that creates the sharp decrease within the production of cotton. So for the study, of interest is that the leaf instead of the entire cotton. About 85%-92% of diseases occurred on the cotton leaves are like Alternaria, Cercospora, Red spot, white spot and Yellow spot on the Leaf

***Index Terms* - Convolution Neural Network, Keras, Tenserflow, MATLAB**

I. INTRODUCTION

India is an agricultural country as per the Observation in India Most of the people depend upon agriculture. Farmers have a good range of multiplicity to pick suitable crops for his or her farm. However, the cultivation of those crops for optimum yield and quality produce is usually technical. The disease diagnosis is restricted by human visual capabilities because most of the primary symptoms are microscopic. This process is tedious, time-consuming. Nowadays within the area of research, a major concern is an identification of the symptoms of the disease by means of image processing. The farmers are struggling during their lifestyle for a way to affect the disease of the cotton leaf. There a requirement for a disease diagnosis system which will support farmers. This technique focuses on disease identification by processing acquired digital images of leaves of the plant.

The many paper we refer they use different algorithms for disease detection of leaf many of use support vector machine (SVM), artificial neural network (ANN) so these all different approaches use to detect disease. The main part or advantage Of our project is we provide solution for disease and give the information which pesticide or insecticides are suitable for that Disease. That can help farmers to stop disease from spreading and crops give better results when we give proper treatment of these crops. The following sample images from our dataset.

II. Litrature Survey

[1] Pranita P.Gulve,Sharayu S.Tambe, Madhu A.Pandey,Mrs S.S. Kanse This paper uses GLCM, thresholding, segmentation, feature extraction for detecting the diseases. They create colour transformation structure for the rgb leaf image , apply colour space transformation structure then image is segmented.Unnecessary parts (green area) within leaf area is removed.Calculate the texture features for the segmented infected object.Extracted features are passed through a pre-trained neural network. Image is taken with the help of digital camera and all images are stored in a JPEG format. Image resize and image filtering is done. Here filtering is important to remove any noise content. For this they used gaussian low pass filter is applied and positive standard deviation sigma. In this the classifier used is euclidian distance classifier.

[2] A. Jenifa, R. Ramalakshmi, V. Ramchandran In this paper uses convolution neural networks, deep neural networks, leaf disease, classification, image processing. This training model contains 500 leaf images and testing model contains 100 leaf images. Cotton leaf diseases to be classified in this are - cercospora, bacterial blight, ascochyta blight, target spot. The tool which they used is MATLAB tool. Diseased cotton leaf taken as input image then input image converted into grey converted image then the noise co-efficient image, then it converted into filtered image, next morphological image then clustered image the segmentation image and finally the output image takes out the diseased part. This system shows the 96% accuracy for the classification of diseased cotton leaves.

[3] Rakesh Chaware, Rohit Karpe, Prithvi Pakhale, Prof. Smita Desai This system identify three diseases- aleternaria atternata, anthracnose and bacterial blight. This system works parallel for healthy and defected leaf image. Two images has Been taken one for the healthy leaf other for the defected leaf. Disease detection starts from training process. In training process resizing of the healthy and defected image has been done. Then convert rgb to grey scale image then apply stem, stairs, canny edge detection, surf, entropy, warp images. This technique is applied on both the samples healthy as well as defected. Once the training process of first phase samples is finished, comparison has been done on the basis of values obtained for all the parameters used.

[4] In this paper they're using Crop Image, Agriculture image Processing, Image segmentation, Histogram Equalization. They divide the diseases within the following three categories: Bacterial disease, fungal diseases, viral disease. The image processing technique is employed for detecting diseases on cotton leaves early and accurately. The processing scheme consists of image acquisition through camera or web, image pre-processing includes image enhancement and image segmentation where the affected and useful areas are segmented, feature extraction and classification. Finally the presence of diseases on the plant leaf are going to be identified. For feature extraction, they are using the K-mean clustering algorithm method for classification and Neural-network as recognizer. The step-by-step procedure is shown as below. 1) RGB image acquisition 2) Pre-processing of image using histogram equalization 3) Resize the image 4) K-mean Algorithm for image segmentation 5) Computing features extraction 6) Classification & Recognition using neural networks 7) Statistical analysis. Study of diseases on the cotton leaf studied by using the image processing toolbox and also the diagnosis Using MATLAB helps to suggest necessary remedies for that disease arising on the leaf of a cotton plant. The accurate recognition for using K- Mean Clustering method the Euclidean distance is 89.56% and the execution time for K-Mean Clustering method using Euclidean distance is 436.95 second and also thresholding is done by a dynamically range [0,1] depending on color intensity from leaves image. Hence that disease detection using K-Mean Clustering method using Euclidean distance is the excellent methods to disease detection on cotton leaves

[5] This system consists of two parts: 1) digital image assessment and feature extraction of sample cotton leaf 2) to implement the back propagation artificial neural network in machine learning. This process has the following five steps- Image acquisition, Image Pre-processing, Image enhancement, Image segmentation and Feature extraction. To classify the quality of cotton leaf diseases, the Artificial Neural Network tool of MATLAB is used. This quality identification is done based on the RGB and HSV components of the image. This ANN tool works on neurons, neurons are further connected to hidden layers of neurons. The ANN tool also has a back-propagation process. The prediction of the outcome is taken randomly by the neural network process. The advantage of this method is it can predict the data correctly with minimized error. This method was able to detect cotton leaf with or without defects from the image.

III. PROPOSE SYSTEM

In a purpose system, we use a real-time dataset that contains various images of cotton disease like bacterial blight, bronze wilt, curly lift, fouler fungal disease. We have given some images for training and some are testing. Initially we take the images from the real time dataset and give it to the model for identifying the cotton disease. As, we give the images to our system, it shows the result in the form of probability. The methodology of cotton disease detection using image processing has the following steps:

1) Image pre-processing:

In these phases, we require better resolution images and with better quality. All these images are resized with specific manner and resolution. These images we remove noise content and rotate the images using a data augmentation process.

2) Image segmentation:

Image segmentation it's a process of dividing a digital image into various sections. Its use to remove the region of the pixel in infected leaf and easily identified the model which part is infected

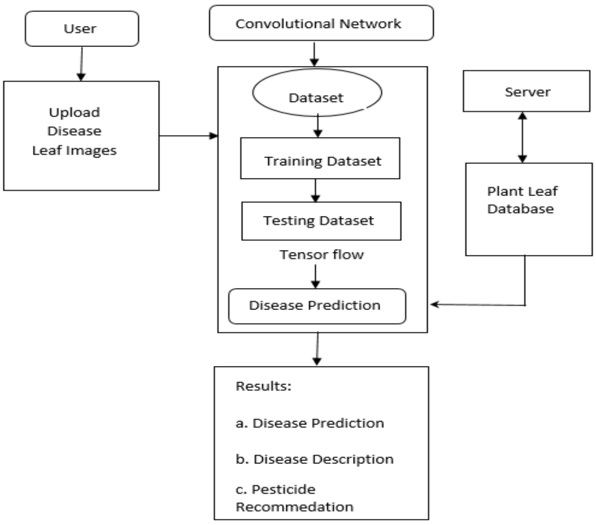

Figure 1: system Architecture

3) Feature extraction:

In this feature extraction process, extract some of the important features of the defected leaf. It can create colored structure and convert the color value from RGB components of defected parts of cotton leaf image. The feature we can use to train our neural network. When all processes are done then we give the train and test data to the model and apply the CNN algorithm. The Following Flowchart you can see how the model is working.

Data Augmentation There are 4 classes of diseases for which the images were collected. The images were less in number. As deep learning based approaches need more images, we have performed data augmentation using MATLAB. This is one time process which gives 10 rotated versions of original image. Therefore lot of data is generated which is helpful to training model. We have used various data science related libraries like keras, tensorflow, sklearn, opencv, matplotlib, numpy etc. For the purpose of building keras model we have used sequential modelling technique. The architecture of model consists of two conv2D layers. Each conv2D layer is followed by activation layer named ‘relu’ and maxpooling layer. Maxpolling is that the highest value they can catch and declare. Once the data is available at final maxpooling layer, it is subjected to set of fully connected neuron as they are in ANN. For this purpose flattening is done and dense layers are added. Flatten is convert output to in one dimensional array that work is done by flatter. This creates the architecture of deep learning model which will be trained using the data which was uploaded earlier.

The data is available on the colab server. Path variable will read the images from the path one by one. Each image is read using opencv library. Subsequently, the images are resized with dyadic image processing. The paths of the images also tell about the Class of each image which is extracted and stored in a variable called label.

The data and label lists are converted into numpy arrays for the purpose of training the model. The data train: test split ratio is 75:25. Runtime data augmentation during training is also made available to the optimizer. For building the model, the classifier is notified that the image dimensions are 128*128*3 along with 4 classes. We have used categorical cross-entropy for measuring the losses during training process which are monitored continuously. We have used Adam optimizer which is latest optimizer that SGD. The training process gives a trained model which can further be used for testing purpose. For testing, we have uploaded the .zip file that contained test images. The performance is checked for images from this .zip file after unzipping it.

IV. Algorithm Used

Traditional feature learning methods rely on semantic labels of images as supervision. They usually assume that the tags are evenly exclusive and thus do not point out towards the complication of labels. The learned features endow explicit semantic relations with words. CNN itself is a technique of classifying images as a part of deep learning. In which we apply single neural network to the full image.

I. Accepts a volume of size $W1 \times H1 \times D1$

II. Requires four hyper parameters:

Number of filters K
Their spatial extent F
The stride S
The amount of zero padding P

III. Produces a volume of size $W2 \times H2 \times D2$ where:

1) $W2= (W1-F+2P)/S+1$
2) $H2= (H1-F+2P)/S+1$ (i.e. width and height are Computed equally by symmetry)
3) $D2=K$

IV. With parameter sharing, it introduces F*F*D1 weights per filter, for a total of (F*F*D1)*K weights and K biases. In the output volume, the dth depth slice (of size W2*H2) is the result of performing a valid convolution of the dth filter over the input volume with a stride of S, and then offset by dth bias.

V. A common setting of the hyper parameters is F=3, S=1, P=1 However, there are common conventions and rules of thumb that motivate these hyper parameters.

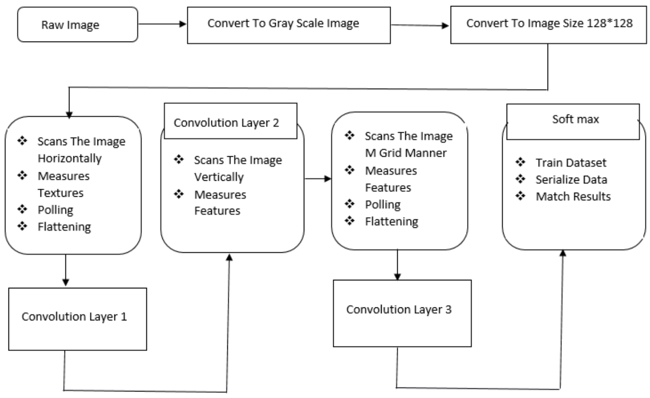

Figure 2: Algorithm Flow

CNN uses the layers for image processing. If the image having the more than one objects then the CNN recognizes the edges and classify the image accordingly. Pixel is the smallest portion of the image. One single image contains number of pixels. These pixels group together it makes entire image. CNN uses feature detector. Feature detector used to detect significant features of image data in order to provide detection. It is the smallest matrix of weights. To reduce the bigger images into smaller images strides are used. Stride is the number of pixel by which we slide our filter matrix over the input matrix. This process is called convolution. By this the shape of the input image is modified feature detection there by detecting the particular feature from the input image and to get the information about that feature. This is called the feature map. Large images takes lot of time. It is easier to process small images in faster manner.

For the image identification we use tensor flow. Keras layer like input layer, dense layer, convolution 2D layer, Maxpool2D layer, activation layer and Flatten. The keras modelling are different techniques but we use sequential modelling. In these sequential modelling we tell that execute the step by step layer and the all network are made. We use LeNeT class to classify the width, height and depth of the image. Softmax classifier can give us probability to the result whether these image are powdery disease or foliar disease or other disease and out of 100% what the probability of the disease they can give us if the probability is 90% so model is good. We can split the path and set the string through labels. The raw pixel intensities to the range to [0, 1] and in grey scale image maximum value are 255 and minimum value is 0. So we divide 255 to 255 minimum value is 1 so it can be normalized. We take the 25% data to test and 75% data to train.

Test Case 1:

Train model with 10 epoch

In test case 1 we perform operation on 10 times of single image the training accuracy of 1st epoch is 0.4815 and valid accuracy of 1st epoch is 0.7624 so our model is not that much good but when we process done in till 10th epoch the training accuracy was 0.8815 and valid accuracy is 0.9424 the result is model was perform good and its give the accuracy of 94% accurate. Here the graph

Figure 3: Model Accuracy Graph

Test Case 2:

Check the model and we give sample image Bronze wilt disease. The model gives the probability which disease this is they can compare all the disease in the dataset and give the percentage the output is Bronze wilt disease - 0.94, Curly Leaf -0.77, Foliar fungal-0.14, Powdery Mildew-0.22, Fresh Leaf 0.61 so the model gives these result and percentage of Bronze wilt Disease is 0.94 so model is working properly. We give this image.

Figure 4: Bronze Wilt Disease Image

The result we give in graph so see this graph for better understand.

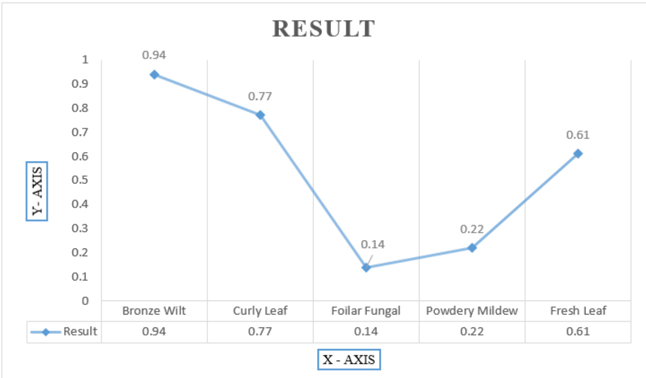

Figure 5: Graph of Bronze wilt

Test Case 3:

Same we give second image of curly leaf and process is same they compare all the disease and give the percentage of all disease the output is Bronze wilt–0.41, Curly Leaf -0.77, Foliar fungal-0.41, Powdery Mildew-0.22, Fresh Leaf 0.61model gives proper output. We give this image.

Figure 6: Curly Leaf Disease Image

The result we give in graph so see this graph for better understand.

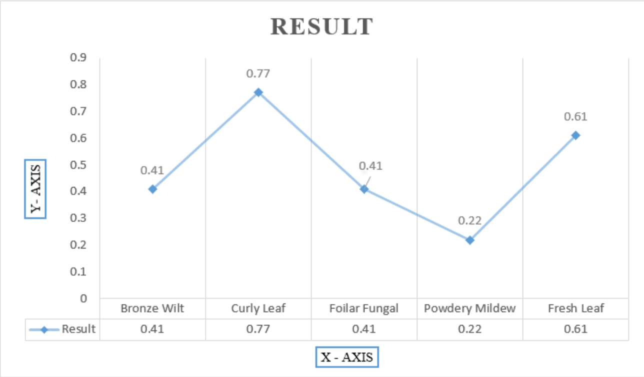

Figure 7: Graph of Curly leaf

The conclusion of these test cases is when we do more operation on image the model can give you accurate output and when we provide original images the model is working well with proper accuracy. We test more on that but the hardware requirement was not fulfill for these operation.

V. Other Recommendations

We took image data set of cotton leaf from the field After giving input leaf image it shows the disease percentage occur on the input leave it will be very beneficial for the farmers to the applied pesticides solution on the particular portion of infected leaf. The dataset of cotton crop was not available on the internet so we went to the cotton field ourselves and took photos of the cotton leaves and made a dataset from it. Images and the same dataset we have used in our system. Solution in the form of pesticides.

VI. SOME COMMON MISTAKES

It will only show the result of cotton leaf disease if we provide other input images rather than cotton leave it will show the result between our define disease classes and this is the only drawback of our system at what time we can provide limited set of images and if we try to provide five or more than five images at one time then system may be hang. There are number of diseases on the cotton plant like seedling diseases bacterial diseases boll roots but our system only predict the disease which are occur on the leaf of cotton like foliar fungal, bronze wilt, mildew, curly leaf etc.

Our system can input a lot of photos at once so it helps to save the farmer's time. And if more than five photos are given for input in our system, our system may hang or the user may have to face some system problems.

VII. Conclusion

The Algorithm will help the end user to segregate the infected crop based on percentage of infection to take preventive measures at as early stage as possible. The algorithm will help in minimizing the use of pesticides thereby improving the environment and ecological balance. The proposed work has vast applications to help the Indian farmers in early identification of cotton crop diseases.

VIII. FUTURE SCOPE

In future we can expand disease dataset for detect more disease. Work on accuracy that can improve user result and give the correct pesticides or insecticides to the user also we built for mobile friendly app so process done on cloud. That one thing can benefit of user because any device they have the process can work.

REFERENCES

[1] Pranitha P. Gulve, Sharayu S. Tambe, Madhu A. Pandey, Mrs S.S.Kanse, “Leaf Disease Detection of Cotton Plant Using Image Processing Technique”, 2015

[2] A.jenifa, R. Ramalakshmi, V. Ramachandran,” Classification of cotton leaf Disease using Multi-Support Vector Machine”, 2019

[3] RakeshChaware, RohitKarpe, PrithviPakhale, Prof.SmitaDesai, "Detection and Recognition of Leaf Disease Using Image Processing", (2017)

[4] Vijay S. Bhong, B.V. Pawar,”Study and Analysis of Cotton Leaf Disease Detection Using Image Processing”, 2016

[5] Nikhil Shah, Sarika Jain, “Detection of disease in Cotton leaf using Artificial Neural Network”, 2019