e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:01/January-2021 Impact Factor- 5.354 www.irjmets.com

# COTTON LEAVES DISEASE DETECTION AND CURE USING DEEP LEARNING

**Rahul Mhatre[*1], Vishal Lanke[*2]**

[*1,2]Department of M.Sc.IT, SIES(Nerul) College of Arts, Science and Commerce, Mumbai, India.

## ABSTRACT

In India, Cotton is considered to be one of the most important cash crops, as many farmers grow cotton in large quantities. Cotton diseases, in the last few decades should lead to significant loss of productivity. Early detection of cotton diseases diagnosis is important. The goal of this research is to introduce a system that uses a convolutional neural network method to automatically diagnose cotton leaf diseases. Separation based on selection of appropriate features such as color, texture of images created using Deep Learning Techniques. The images are acquired from various online sources.

## I. INTRODUCTION

Agriculture has played a major role in the financial development of most agricultural countries including India. The cotton crop is mainly found in North India. There are a variety of diseases that have had a negative impact on the performance of cotton plants over the past few years. The most important requirement is the acceptance of the infection and its planning.

Plant diseases need careful diagnosis and timely management to protect the plants from severe losses. It is a well-known fact that the disease is diagnosed by farmers too late and is very difficult and expensive to control. The main reason for this is the inaccessibility of experts close to farmers because a large number of professionals are difficult to produce in the country, so the availability of a Computer based expert program can be a blessing for farmers.

The task of this project is to create a system that is available to farmers through the transfer of images to a centralized specialist system to obtain information on the disease and its remedies. In this way, human specialists will provide diagnostic technology to the farmers. The computer scientist will use this information to create a training set that will be used on images to ensure that the disease can be detected in sufficient proportions. The pattern matching algorithm will be developed for the purpose of initial detection of the disease with acceptable accuracy levels at a fast pace.

The diagnosis of cotton leaf disease is the process of diagnosing diseases by analyzing their physical properties. The process of removing visible structures from images is known as the feature extraction. The feature extraction process can be performed using various feature descriptors. Feature descriptors are then transferred to a separator to check the feature. Separator is an algorithm, which is used to classify an element on the basis of its similarity to the training database. The training database is a collection of features that were previously excluded from known objects. Diseased leaves are classified on the basis of their similarity to the training databases of disease samples previously described by the descriptors.

Our goal is to solve the problem of getting cotton diseases using automatic image processing techniques from the input image. The diagnosis of the disease will depend largely on the presence of the disease in the cotton leaves, which can be used to diagnose using a separator.

There are some problems with field crops such as identifying malnutrition in plants, identifying various diseases, various pests affecting crops. Each issue has its own significance. Among these, one issue is the detection of pests so that appropriate measures can be taken to control them leading to reduced losses. In the event of any such situation farmers know about the pest, then they can take appropriate action and control the situation but if the farmers do not have the right information, then misidentification of any pests can occur and improper control equals not to touch pesticides can be used leading to waste of hard work and money. Most importantly it can lead to serious plant problems.

The classification using the combination method has many limitations when calculating the number of collections available. The leaf recognition algorithm using the feature recognition system is used. The identification of plant leaves is done using the Convolutional Neural Network. Features are extracted and processed by PCA to enhance those features to include them on CNN. A new method of image recognition based on Wavelet Transform and Singular Value Decomposition that is capable of retrieving multiple images such as a target image. By using the image of wavelets in the upper bag is extracted and enhanced edge details further rotation is used, which provides better improved boundary details with wavelet features such as strength and entropy. The wavelet features and methods are integrated into our system to provide better accuracy results. Also PCA analysis is used to reduce the factor which reduces the time required for processing. Progressively these diseases are classified using the back convolutional neural network. This method reduces processing time and accuracy is increased.

## II. LITERATURE REVIEW

According to **"Ole Mathis Opstad Kruse et al."** (2014), they examined the effectiveness of the four-dimensional method of identifying pixels representing damaged areas in leaf areas. Consistent with the MIA method of the T2-linked pattern, RSS statistics and partial analysis of the monitored methods of use were also applied to the K-means clustering method. This method is significantly superior to the other three methods of separation in pixel detection with the highest accuracy, accuracy, real positive and F-score ratio and the lowest false positive and time calculation. A real good rate of 80% indicates that 20% of injured pixels are falsely identified as healthy.

According to **"Mahesh Shivaji Dange et.al."** Infected leaf characteristics are compared to normal leaf texture features. Improved processing has four main steps, first the RGB color image creation is created, and then the RGB image is converted to HIS because RGB is for color reproduction and HIS for color definition. Then the green pixels are hidden and removed using a certain amount of limit, then the image is split and the useful parts are extracted, finally the texture statements are calculated using spatial Gray Level Dependence Matrices.

According to **"Tushar J.Haware**"( 2015) an algorithm that works well with high precision mixing is used to detect plant diseases. K-means integration is used for partitioning and the method of computational computing. The red-colored, red-blue Zeros pixels and pixels on the boundary of the infected group are removed.

**"P. Revathi et.al." (2012),** developed an advanced computer technology to help the farmer make a higher decision regarding many aspects of the harvesting process. Proper diagnosis of Foliar plant diseases is very important for increased productivity. Technological advancements have been developed using cellular photographic signals of cotton dots and disease classification using the neural network. The classifier is trained to perform clever farming, including early detection of diseases in the fields, to apply fungicide etc. The work is based on the techniques of image classification where photographic images are processed to enrich first. Thereafter the color rendering of R, G, B to identify the target areas of disease. Image elements such as border, shape, color and texture are removed from disease areas to detect disease and to control insect recommendations.

**"A. Camargo et.al. " (2009).** described a photographic-based approach that identifies the obvious symptoms of plant diseases, starting with the analysis of color images. The processing algorithm starts by converting the RGB image of a plant or a diseased leaf into H, I3a and I3b color conversion. I3a and I3b mutations are developed from the original color conversion I1I2I3 to meet the data needs of plant diseases. The converted image is then separated by analyzing the power distribution in the histogram. Instead of using the traditional method of selecting the minimum area as a boundary, a set of local maximum is obtained and the cut-off maximum value is determined according to their position in the histogram. This method is especially useful when the target in the image data set is a large power allocation. In the experiment, when the image was separated, the extracted region was rescheduled to remove pixel regions that could be considered part of the target region. The process is accomplished by analyzing the location of each pixel and the transition gradient between them. To test the accuracy of the algorithm, hand-separated images are compared to those automatically separated. The results showed that the advanced algorithm was able to identify a sick region even if that region was represented by a wide range.

## III. RESEARCH METHODOLOGY

The whole process of training a model for plant disease detection using CNN in-depth is described in detail. The complete process is divided into several required sections in the steps below:

**Dataset**:

Appropriate data sets are required for all phases of object recognition research, from the training phase to testing the effectiveness of visual algorithms. All data collected from the database were downloaded from the Internet, searched for diseases and plant names from various sources. The dataset will be classified into diseased leaves and healthy leaves. The dataset will be trained by using a Deep Learning model (Convolutional Neural Network). Therefore, a deep neural network can be trained to separate adjacent leaves.

The next step was to enrich the database with additional images. The main purpose of the research presented to train the network is to study the features that distinguish one category from the other. Therefore, when you use multiple additional images, the network's ability to read relevant features is increased.

**Image Pre-processing and Labelling:**

In order to get a better feature output, the final images intended for use as a deep neural network configuration database will be processed to determine consistency. In addition, the process of image editing involves cutting all the images using machine learning, making a square around the leaves, to highlight the area of interest (cotton plant leaves). During the data collection phase, images with minimal editing and size less than 500px are not considered valid image data. Additionally, only images where the interest region has the highest resolution are marked as eligible for the dataset. It was ensured that the pictures contained all the details needed to learn the feature. Dataset images were scaled to reduce training time, which was automatically calculated by text in Python, using the Open CV framework.

Many resources can be found by searching across the Internet, but their suitability is often unreliable. In an effort to ensure the accuracy of the classrooms in the database, which was initially collected by keyword search, agricultural experts examined leaf images and labeled all images with the appropriate stem of the disease. As is well known, it is important to use accurately classified images for training and validation data. Only then can a suitable and reliable acquisition model be developed. At this stage, duplicate images that were left after the initial collection and collection of images were removed from the database.

**Augmentation Process:**

The main purpose of using the augmentation is to enlarge the dataset and to bring less distortion to the images that help to reduce excess during the training phase. In machine learning, as well as in mathematics, override occurs when a mathematical model describes a random sound or error rather than a basic relationship. Image enlargement contains one of several transformation modes that include affine modification, modification, and simple image rotations. Positive variables have been used to express translation and exchange (corresponding conversions and adding vector, resp.) When all the same lines in the first image are still the same as the output image. To obtain a transformation matrix, three points from the first image were required and their corresponding locations in the output image. For a change of perspective, a 3 × 3 matrix for change was required. Straight lines can stay straight even after conversion. With the add-on process, simple image rotation is used, as well as rotation on different axis at different levels. In this section, to make the process of adding multiple images from the dataset, a specific application built into Python using the OpenCV library, may change the parameters of the change during startup, which improves flexibility.

**Neural Network Training:**

It is proposed to train a deep convolutional neural network to model image separation from the dataset. There are several in-depth study frameworks, such as the Python Theano Library and the Lua Extensive Electronic Library, Torch7. In addition, there is Caffe, an in-depth source learning framework developed by BVLC that contains a CaffeNet reference model. CaffeNet is a deep CNN with multiple layers that continuously integrates features from input images. Specifically, the network consists of eight reading layers and five specification versions and three fully integrated layers. The construction of CaffeNet is considered a start, but has been redesigned to support our 15 sectors. The final layer was changed and the softmax layer extraction was made a parameter according to the needs of the presented lesson. The convolutional layer is an important basis for building a convolutional neural network. Layer parameters contain a set of readable characters that have a small reception field but extend to the full depth of the input volume. Each convolutional layer contains maps of equal size, $Mx$ and and, with a kernel of size $Kx$, and $Ky$ is moved to a specific region of the input image. The $Sx$ and $Sy$

sketch elements define how many pixels the filter / kernel skips on $x$- and $y$- pathways between subsequent interactions. The size of the output map can be defined as:

$$M_x^n = \frac{M_x^{n-1} - K_x^n}{S_x^n + 1} + 1,$$

$$M_y^n = \frac{M_y^{n-1} - K_y^n}{S_y^n + 1} + 1,$$

where $n$ indicates the layer. Each map in layer $Ln$ is connected to most $Mn-1$ maps in layer $Ln-1$.

Rectified Linear Units (ReLU) are defined as:

$$f(z_i) = \max(0, z_i)$$

Deep CNN with ReLUs trains several times faster. This method is used for the extraction of the entire convolutional and fully connected layer. Apart from the output, standard installation is not required; is used after ReLU disconnection after the first and second convolutional layers because it lowers the top-1 and top-5 values. At CNN, neurons inside the latent layer are separated into "feature maps." The neurons within the feature map share the same weight and bias. Neurons within a map element search for the same feature. These neurons are different because they are connected to different neurons in the lower layer. So in the first hidden layer, the neurons within the feature map will be linked to different regions of the input image. The hidden layer is separated by feature maps where each neuron in the feature map looks at the same element but at different image capture locations. Basically, a feature map is the result of applying convolution to an entire image. The features of each layer are displayed in a separate block, where visibility represents the strongest performance of a given feature map, from the first convolutional layer, where the elements from individual pixels to the simple lines, in the fifth convolutional layer where layered features and parts of the leaves are displayed.

Another important layer of CNN is the integration layer, which is a type of offline reduction. The functionality of the pools provides a kind of translation flexibility; it works independently across the input depth slide and makes its size geographically. Excessive blending is used to benefit the reduction of excessive excess. And in favor of reducing overeating, the quit layer is applied to the first two layers that are fully connected. But the failure to drop out of school is that it increases training time by 2-3 times compared to a normal neural network of direct construction. Bazesian performance tests have also proven that ReLUs and school dropouts have the effects of collaboration, which means that they are beneficial when used together. CNN's advancement refers to their ability to study medium image representation with contradictory and low-level features designed for alternative image classification methods.

**Testing**:

A common way to measure the performance of artificial neural networks that separate data from the training set and the test set and then train the neural network in the training set and use the predictive test set. Therefore, as the initial results of the test set and our predicted model of results are known, the accuracy of our prediction can be calculated. A separate experiment is performed with the original image database, in which a larger image database is trained.

Accuracy test, used a 10-fold cross verification process to test the prediction model. The process of cross-verification was repeated after every 1,000 training sessions. The full measurement result of the test is symbolically represented as a top-1, to test whether the upper class (with the highest probability) is the same as the target label. The top-5 error level exists to test whether the target label is one of the top 5 predictions, the ones with the highest five possibilities The number of images used for the verification test from each category labeled is given in Table 1. Test results are presented in Section 4, for both complete databases and for each class separately.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:01/January-2021 Impact Factor- 5.354 www.irjmets.com

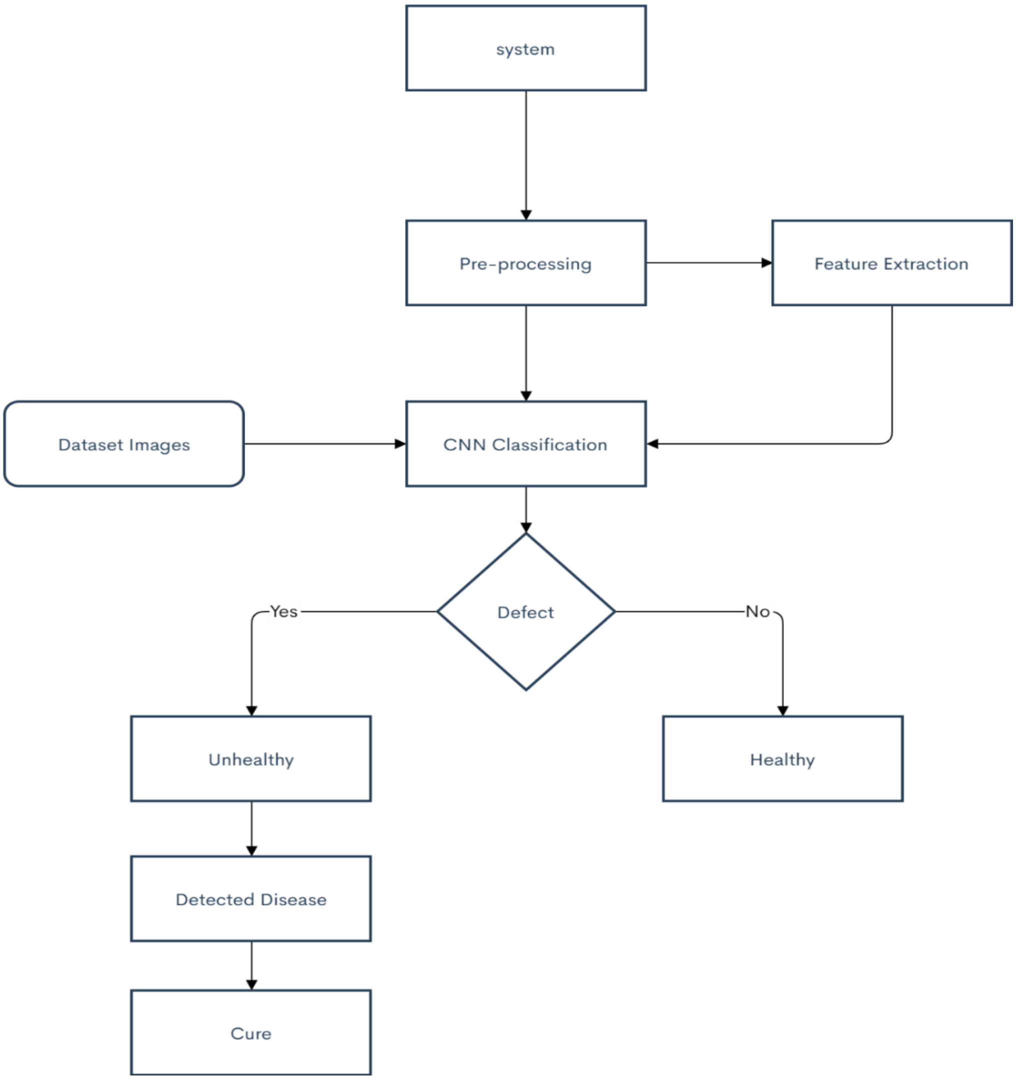
**Fine-Tuning:**

Good adjustment aims to increase the efficiency or effectiveness of a process or function by making small changes to improve or maximize the effect. The split function in the original CaffeNet model is a softmax classifier that combines a number of class images of ImageNet dataset. A well-prepared test requires a little learning, but much faster than learning from scratch. To begin the fine-tuning process, this softmax separator is removed, and a new one is introduced at random values. The process of performing finetuning was repeated to change the layers of the hidden layers and hyperparameters. The most appropriate model for the detection of plant diseases was obtained by the process of adjusting the boundary tests.

**Deployment:**

This step includes the deployment of the trained model by using flask. Flask is a small web framework written in Python. Divided as a small task because it does not require any special tools or libraries. It does not have a database extraction layer, form verification, or other items where existing third-party libraries provide similar functions. However, Flask supports extensions that can add app features as if they were made to Flask itself.

## IV. BLOCK DIAGRAMS




## V. CONCLUSION AND FUTURE WORK

The proposed project subsists by collecting the necessary data for various cotton diseases by examining different sectors. Image processing techniques and classification techniques are used to identify cotton leaf

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:01/January-2021 Impact Factor- 5.354 www.irjmets.com

diseases. Factors such as color, shape and texture are useful in pattern recognition, classification, free accuracy and errors are calculated. Future work will be to build an efficient, robust detection system for automatic detection of various plant diseases. Classifier will be based on a selection of different features or a combination of different algorithms for fast diagnostic tests.

**Images of machine learning disease detection:**

(a)

Normal Leaf (b) Red Spot (c) White Spot (d) Leaf Crumple

Cotton Leaves (a) Normal (b)(c)(d) Diseased

## VI. REFERENCE

[1] V. A. Gulhane and A. A. Gurjar, "Detection of diseases on cotton leaves and its possible diagnosis," International Journal of Image Processing (IJIP), vol. 5, no. 5, pp. 590–598, 2011.

[2] P. Revathi and M. Hemalatha, "Classification of cotton leaf spot diseases using image processing edge detection techniques," in Emerging Trends in Science, Engineering and Technology (INCOSET), 2012 International Conference on. IEEE, 2012, pp. 169–173.

[3] A. A. Sarangdhar and V. Pawar, "Machine learning regression technique for cotton leaf disease detection and controlling using iot," in Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of, vol. 2. IEEE, 2017, pp. 449–454.

[4] A. Parikh, M. S. Raval, C. Parmar, and S. Chaudhary, "Disease detection and severity estimation in cotton plant from unconstrained images," in Data Science and Advanced Analytics (DSAA), 2016 IEEE International Conference on. IEEE, 2016, pp. 594–601.

[5] P. Rothe and R. Kshirsagar, "Adaptive neuro-fuzzy inference system for recognition of cotton leaf diseases," in Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), 2014 Innovative Applications of. IEEE, 2014, pp. 12–17.