ISSN No:-2456-2165

Mulberry Leaf Disease Detection

Nagaveni B Nimbal[1], Namratha S.[2], Arshitha K.[3], Chaithra Shetty S.[4], Bhuvanashri KB[5]
Asst. Professor[1], Student[2,3,4,5]
Department of Computer Science
KSSEMKSSEM Bangalore, India

Abstract:- Sericulture is an important domestic industry. In India,it is one of the eco-friendly industries. India is the only country where all five recognized commercial silks are made, namely mulberry, tropical tasar, oak tasar, eri and muga. Sericulture is labor intensive, providing jobs in India for over 8 million people, and serving Indians as a tremendous source of revenue. Silkworm is sericultural foundation. Commercial silk is developed through the production of different types of silkworms, of which BOMBYX MORI, originally from Asia, is the most widely and economically used, Mulberry is significant sole nourishment for mulberry silkworm, which exclusively benefits from the leaves of mulberry plant. These silkworms are totally domesticated and reared indoors. Other uses of mulberry leaves are seen in the fields of health and skin care. These mulberry plants include a high pace of yield disappointment and are over the top expensive for creation, so should be dealt quite well.Our goal is to overcome these problems using a farmer-friendly system where the result involves cure of the disease and the fertilizer or pesticide proportion to be used are displayed on the user interface.

Keywords:- Plant disease, Image preprocessing, disease classification, EHD, Support vector method, KNN, DT.

I. INTRODUCTION

Sericulture is an agro-based industry. India is second largest producer of silk. Sericulture has following highlights: high work potential, gives energy to town economies, Low Gestation and High Returns, women friendly occupation, Eco-accommodating Activity. India has the one amongst a form qualification of being the main nation delivering all the five known business silks, especially mulberry. Mulberry silk is more produced in India from Karnataka, Andhra Pradesh, Tamil Nadu, Jammu & Kashmir and West Bengal, whilethe non-mulberry silks are produced in the states of Jharkhand, Chhattisgarh, Orissa and north-eastern states.Agriculture is a key source of livelihood. Agriculture provides employment opportunities for village people on large scale in developing country like India. Most of Indian farmers are adopting manual cultivation due to lagging of technical knowledge, when plants are affected by diseases through leaves that will affect production of agriculture and profitable loss.

Leaves are important for fast growing of plant and to increase production of crops. Identifying these diseases in plants leaves is challenging for farmers. They are integral part of the plant system they play a major role in preparation of food in plants. The detection process is essential one in the agriculture field in the beginning of a series of activities to fight the plant disease and reduce their spread as soon as they appear on leaves of the plant. Technology has grown to such an extent that a machine is capable enough to predict the leaf disease without human intervention.to predict these diseases we use KNN and SVM algorithm. Mulberry silk originates from the silkworm, Bombyx mori L7 which exclusively benefits from the leaves of mulberry plant. These silkworms are totally domesticated and reared indoors. Other uses of mulberry leaves are seen in the fields of health and skin care.There is an option for consulting experts, but farmers living in remote villages may not be able to afford travelling expenses because he laboratory test results could take a few days and farmers have to travel many times and government experts visit the farm a few villages, but not regularly.

There is a possibility for counselling specialists, yet farmers living in remote towns probably won’t manage cost of travelling expenses because the test results from the lab could few days and farmer needs to travel ordinarily and in hardly any towns specialists from government division visits the sector however not on standard premise.

II. RELATED WORK

The image of the leaf is taken and then converted into greyscale image. Using texture feature including contrast, local homogeneity, cluster shades are also used. For image classification, SVM classifier and minimum distance criterion is used. The data collection is done by taking pictures and then image is preprocessed using image annotation and augmentation. Image analysis is done using multiple extractors. For experimentation dataset is divided into testing, training and validation sets. System can classify the diseases using algorithm. the image segmentation including all existing factors using the HSI color system, here H component is used to segment spots and to reduce illumination. the affected region and leaf segment area respectively. In the final step, the disease is classified using quotient calculation of leaf and lesion area. According to this research, this method is fast and calculation of leaf disease severity is accurate, here leaf area is calculated by using threshold segmentation disease detection using KNN and SVM algorithm. Feature extraction and statistical features. Different statistical features energy, sum entropy, covariance, information measure, entropy contrast. The system can detect the disease with accuracy.

III. MULBERRY DISEASES

A. Powdery Mildew

Powdery Mildew Disease of mulberry is caused by fungal Pathogen, *Phyllactiniacorylea*. The major symptom of this disease is an appearance of white powdery patches on the lower surface of the leaves. When the disease is severe, the white powdery patches turn in to brownish-black; the leaves become yellow, coarse and lose their nutritive value. The severely infected leaves will be become powder brittle even after gently crumple by hand.

Fig. 1: Leaf spot

B. Leaf Spot Disease

Leaf Spot Disease of mulberry is caused by a fungus, *Cercosporamoricola*.

At the beginning of the disease there will be small light brown irregular spots appear on mulberry leaf surface. Later, these spots enlarge and join together leaving with characteristic 'shot hole' and yellow patches around the brownish spot and wither off.

Fig. 2: Leaf spot

C. Leaf rust disease

Mulberry leaf rust disease is caused by *Peridiospora mori* fungus.At the rust disease initial stage, the symptoms is a circular pin head sized light brown spots and becomes darkish brown spots as disease advances. The leaf loss is up to 5 %. Rust severity significantly increased with increasing shoot age, irrespective of pruning time.

Fig. 3: Leaf rust

D. Leaf curl

Curly leaves of mulberry revealed the presence of Taeniothrips, attacking mulberry leaves. The attack of this taeniothrips to mulberry leaves is known from India and Sri Lanka. They injure epidermal tissue and affected leaves show early maturity, depletion of moisture, reduction in crude protein silkworm rearing. Taeniothrips affected leaves generally show streaks in the early stage of attack whereas blotches are observed at the advance stage and ultimately become yellowish-brown on maturity.

Fig 4: Leaf curl

IV. LITERATURE REVIEW

Detection of plant leaf diseases using image segmentation and soft computing techniques. Information Processing in Agriculture.

A generic algorithm for image segmentation is presented. Soft computing techniques are used for image segmentation. Here, the images are converted into greyscale, and texture featuring methods such as local homogeneity, cluster shades are also used. SVM classifier and minimum distance criterion are used in image classification.

A. Plant Disease Detection in Image Processing Using MATLAB

Plant leaf disease detection is done using KNN and SVM algorithms. The different statistical features energy, sum entropy, covariance, information measure, entropy contrast is being considere.

B. Feature Extraction

In this feature extraction the shape, colour& texture is identified. The aim of this phase is to extract features such as colour and shape. Two shape features such as area and perimeter are extracted This includes two phases: image resize and image filtering.

C. Classification

The final stage of the work is selection of suitable classification algorithm for classification of leaf disease to the category they belong.Support Vector Machine (SVM) is a supervised machine learning algorithm that can be used for classification. SVM is popularly used in texture and classification. After extracting colour and texture features of leaves the classification.

V. THE PROPOSED METHOD

A device is being built that automatically identifies leaf disease using image recognition. This in turn helps the farmers in the identifying the diseases at an early stage and provide useful information to control it. This system does few image processing techniques like image acquisition, image pre-processing, feature extraction and classification. Modern agricultural practices assure great development of cultivation.We have other smart agriculture technology models. The Project aims at classifying the diseases in plant leaf at the earlier stage prevent the plants from the diseases. The model is based on Supervised Learning Algorithms to detect the disease which is more accurate for classification purposes. SVM, KNN, DT techniques are used in classifying the plant leaves into healthy or diseased, if it is a diseased plant leaf these algorithms will give the name of the particular disease and control measures. At first an input image is selected from the folder which under goes image acquisition, pre-processing where it undergoes rescaling, converting into grey scale and removal of noise. Then the training images are subjected to feature extraction. The texture features extracted are used for classification. The classification results detect and categorized based on the disease of leaf. The output is obtained at last which indicates the disease name along with its control measures for the better management of the crop.For feature extraction we have used EHD, HOG and GLCM. For classification we have used SVM, KNN and decision tree, where we have applied cross validation for SVM.After the image under goes Acquisition, Pre-processing, Feature Extraction and Classification the result will be obtained in the user interface. This contains cure of the disease and the fertilizer or pesticide proportion to be used and also future preventive measures which helps the farmer to manage the disease.

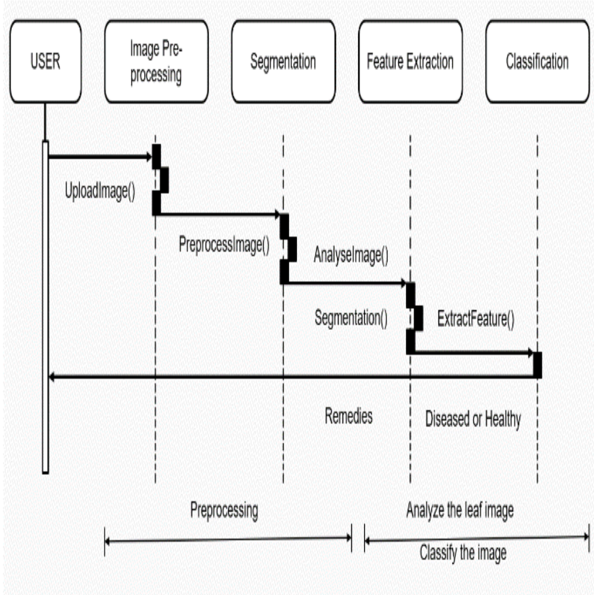


Fig. 5: Design of Application

A. Disease identification is done using following steps:

- *Image Acquisition.*
- *Image Pre-processing*
- *Feature Extraction*
- *Image Classification*

- *Image Acquisition:*

Image acquisition can be definedas the retrieval of an image from some source,typically a hardware-based source, so that it can be passed through whatever processes need to occur later. Performing image retrieval in image processing is the first step of the workflow sequence, because processing without a image is not achievable. The captured picture is totally unprocessed and is the outcome of whatever hardware was used to produce it, which can be very useful for providing a stable reference on which to operate on certain areas. The images obtained using the digital camera is subjected to further pre-processing**.**

- *Image Pre-processing:*

The data collected are typically noisy and come from different sources. In order to feed this collected data into the machine learning model, it should be in standard format or noise free. More often, preprocessing is used to perform steps that minimize complexity and improve the accuracy of the algorithm implemented. Pre-processing is meant to enhance image data by removing unnecessary artifacts or enhance other image features required for further processing. To increase the image quality, the photographs collected from the camera are subjected to preprocessing.

- **Resize:** Some images caught by camera and fed to our AI algorithm fluctuate in size therefore we should establish a base size for all images fed into our AI algorithms.
- **Grayscale:** To store a single colour pixel of an RGB colour image we will need 8*3 = 24 bits (8 bit for each colour component), but when we convert an RGB image to grayscale image, just 8 bit is required to store a single pixel of the image. So, we will require 33 % less memory to store grayscale image than to store an RGB image. It is

ISSN No:-2456-2165

also easier to recognize features of an image when we manage with a single layered image.

- **Noise removal:** Noise is a random variation of image Intensity and visible as a part of grains in the image. Here, we use median filtering to remove the noise. Median filtering is the nonlinear process which is useful in reducing impulsive. The median filter is broadly utilized in computerized image processing just because it preserves edge properties unlike Gaussian blur technique.

➢ *Feature Extraction:*

Feature extraction is the name for methods that select and /or combine variables into features, effectively reducing the amount of data that must be processed, while still precisely and completely describing the original data set.

- **EHD:** The edge histogram descriptor (EHD) is one of the widely used methods for shape detection. It basically represents the relative frequency of occurrence of each local area of 5 types of edges, called a sub-image or image block. The sub image is characterized by non-overlapping blocks dividing the image space into 4x4. So, the partition of image certainly makes 16 equivalents estimated regardless of the size of the original image. To define the features of the image block, histogram is generated for edge distribution for every image block.

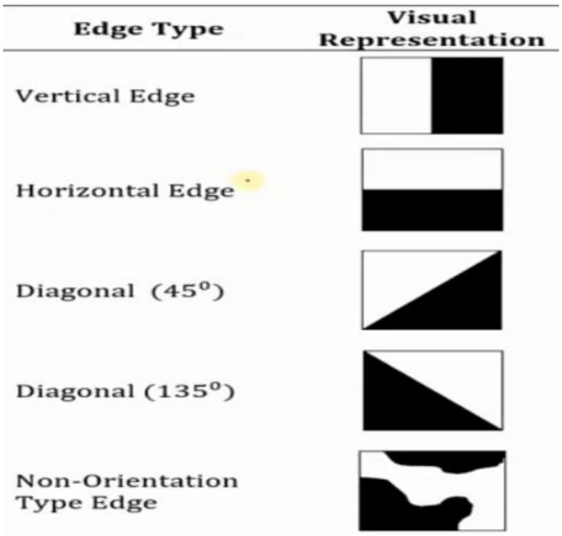


Fig. 6: Five Types of Edges in EHD

The complete semantics for the EHD of 80 histogram bins are listed in Table 1. Each histogram bin meaning should of course be standardized and quantified. For the normalization process, the amount of edge occurrences inside the sub-image for each bin is separated by the total number of image-blocks. The image-block can be a simple device to collect details about the face. This implies that for increasing picture object, it decides if there is at least one edge and which edge is prevalent. When there is an edge, the predominant type of edge between all the 5 edge segments is determined as well. The histogram size of the corresponding edge bin would then increase by one. Then the picture-block does not include an edge for the monotonous area inside the picture.

Histogram bins	Semantics
Bin Counts[0]	Vertical edge of sub-image at (0,0)
Bin Counts[1]	Horizontal edge of sub-image at (0,0)
Bin Counts[2]	45-degree edge of sub-image at (0,0)
Bin Counts[3]	135-degree edge of sub-image at (0,0)
Bin Counts[4]	Non-directional edge of sub-image at (0,0)
Bin Counts[5]	Vertical edge of sub-image at (0,1)
:	:
Bin Counts[74]	Non-directional edge of sub-image at (3,2)
Bin Counts[75]	Vertical edge of sub-image at (3,3)
Bin Counts[76]	Horizontal edge of sub-image at (3,3)
Bin Counts[77]	45-degree edge of sub-image at (3,3)
Bin Counts[78]	135-degree edge of sub-image at (3,3)
Bin Counts[79]	Non-directional edge of sub-image at (3,3)

Table 1: Semantics of local edge bins

➢ *Image Classification:*

Image classification refers to image labelling in one of several predefined classes. Classification of images refers to the task of extracting classes of information from a multiband raster image. The resulting image classification raster can be used to create themed maps. There are two forms of classification, based on the relationship between the analyst and the machine during classification: supervised and non-supervised.Classification of images focuses on the identification of images in one of several classification techniques.

- **SVM:** Here we introduce multiclass Classification model with one-against-all method where there would be a binary SVM to distinguish members of that class from members of that group for each class. It builds templates for k SVM where k is the number of groups. For each of the examples in the i^{th} sequence, the i^{th} SVM is equipped with positive labels, and all other examples with negative labels remain. Thus, given l training data $(x_1,y_1),\ldots,(x_l,y_l)$, where $xi \in R^n$, $i=1,\ldots,l$, and $y_i \in \{1,\ldots,k\}$, the x_i class, the i^{th} SVM solves the following problem:

$$\begin{aligned} \min_{w^i,b^i,\xi^i} \quad & \frac{1}{2}(w^i)^T w^i + C\sum_{j=1}^{l}\xi_j^i (w^i)^T \\ & (w^i)^T\phi(x_j)+b^i \geq 1-\xi_j^i, \quad \text{if } y_j = i \\ & (w^i)^T\phi(x_j)+b^i \leq -1+\xi_j^i, \quad \text{if } y_j \neq i \\ & \xi_j^i \geq 0, \quad j=1,\ldots,l \quad (1) \end{aligned}$$

The penalty variable is where the xi training data is converted to a multidimensional space by function ϕ and function C. Minimizing (1/2)(w)T wi implies we want to optimize 2/||wi||, the difference between two data groups. It is not possible to segregate the data linearly, there is a penalty term C Σlj=1 εij that can reduce the number of training errors. The core idea underneath Algorithm is to find a stability between the terms (1/2)(w)T wiand the training errors.There are functions for decision k after solving the above equation (1).

ISSN No:-2456-2165

$$(w^1)^T\phi(x) + b^1$$
$$\vdots$$
$$(w^k)^T\phi(x) + b^k.$$

We say that x falls within the class with the greatest value of the decision function.

$$\text{class of } x \equiv \arg\max_{i=1,\ldots,k}((w^i)^T\phi(x) + b^i). \qquad (2)$$

- **KNN:** The k nearest neighbor algorithm (k-NN) is a non-parametric approach that is used for classification and regression. The last performance in the KNN group is a time association. The item is listed by plurality vote of its neighbors. Consider the object that is close to the k value, where k represents a positive integer. We have implemented Euclidean, Minkowski, Chebychev, Cityblock, Cosine distances formulas.

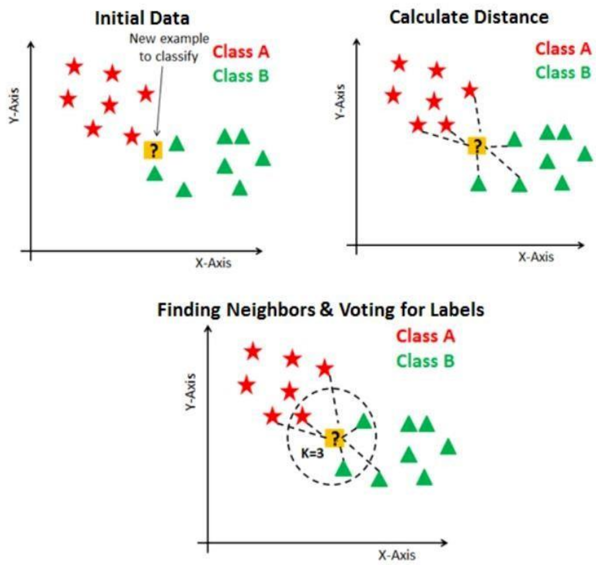


Fig. 7: Working of KNN

The execution of a KNN classifier is essentially dictated by K's decision and also the connected division metric. The gage is determined by the affectability of region measure K assessment, on the grounds that the sweep of the surrounding locale is defined by the division of the nearest neighbor to the query from the Kth and broad K yields distinctive restricted class likelihood. Through introducing the deviations from various groups, an exhaustive estimate of K essentially renders the gage over smoothing and the scheme implementation debases.The estimation of k-NN is perhaps the simplest of all measurements in machine learning.

- **Decision Tree:** The decision tree algorithm is the supervised learning process. It's being used to solve both regression and classification issues. It uses tree representation to deal with the problem. One class represents the leaf node. Initially we consider the whole training set to be the root. Recursive distributions are made on the basis of attribute values. Information Gain used in each level for the identification of the root node attribute. To accurately define gaining information, we need to define a measure commonly used in entropy, the information theory. The classification of test data is done through traversing the tree. Test is applied at the root node to decide which child node the record will come up next. This process is repeated until the record reaches at leaf node. Each leaf node has a unique path from root to root. The path is a rule used to sort the record.

B. Disease Management

After the image under goes acquisition, pre-processing, feature extraction and classification the result will be obtained in the user interface. This contains cure of the disease and the fertilizer or pesticide proportion to be used and also future preventive measures which helps the farmer to manage the disease.

VI. OUTPUT

The implementation of our project is done in MATLAB. We have basically concentrated on 6 types of mulberry leaf diseases. All these data sets were captured by us through mobile camera. We have used 600 images where 420 are used for training set and 180 are used as testing set. At first the image is selected from the folder and after processing happens in MATLAB the output is obtained which indicates the disease of the leaf along with its control measures for the better management of the crop.

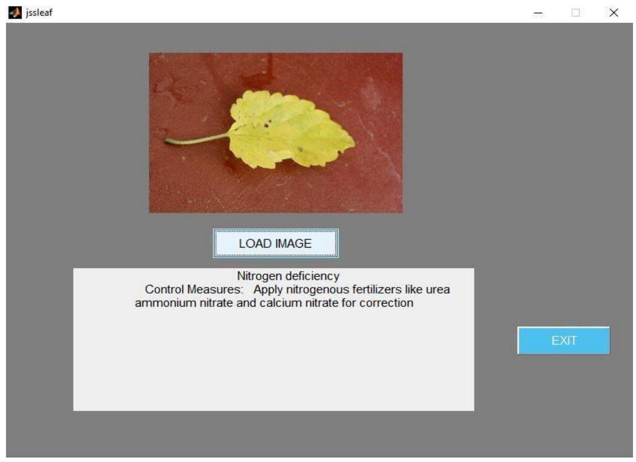


Fig. 8: Nitrogen Deficiency

Figure 8 represents Nitrogen Deficiency. Nitrogen deficiency leaf shows Gradual and sluggish plant development with lower cell growth / vigoration. The young green leaves display chlorosis, the stem is thin and yellowish green and the root development stunted.

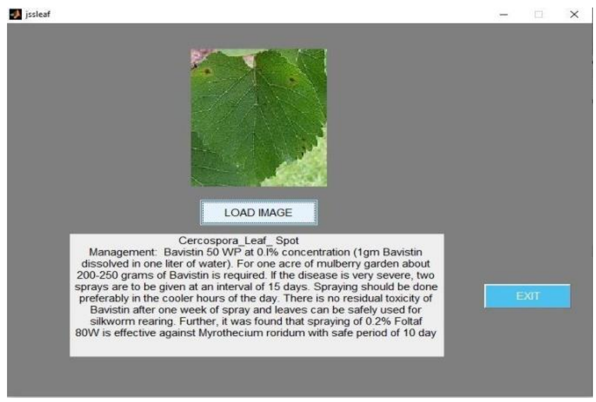


Fig. 9: Cercospora Leaf Spot

Figure 9 represents Cercospora Leaf Spot.Based on the weather and variety, it decreases leaf yields by around 10 – 20 percent. Symptoms are in beginning phases, they occur as little light brown, small dots mostly on stems. As when the disease worsens, the spots are formed to enlarge, coalesce, and shot-holes. Severely infected leaves grow yellowish and prematurely drop down.

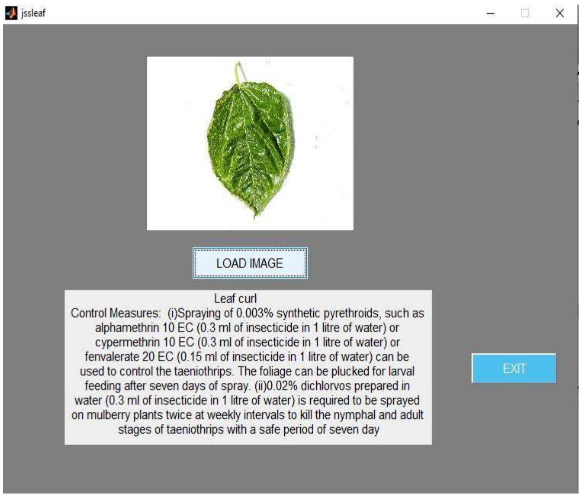


Fig. 10: Leaf Curl

Figure 10 represents Leaf Curl. Curly Mulberry leaves showed the existence of Taeniothrips, assaulting leaves of mulberry. The assault on mulberry leaves by these taeniothrips is reported from India and Sri Lanka. They injure epidermal tissue and significantly impacted leaves show early maturity, humidity depletion, decrease in the rearing of crude protein silkworms.

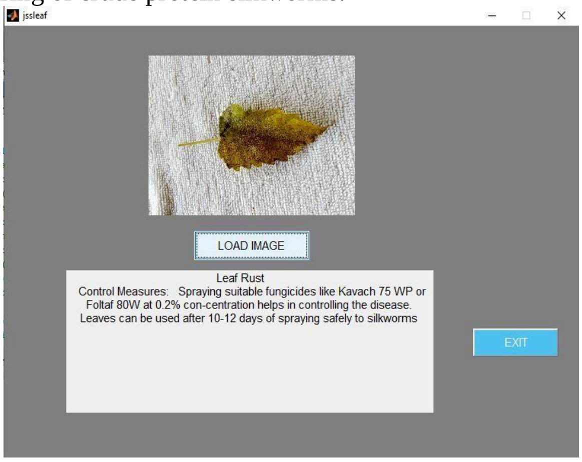


Fig 11: Leaf Rust

Figure 11 represents Leaf rust. Ceroteliumfici triggers herb rust infection. Under the genus Irnperfect fungi, the pathogen suborder Uredinales and family Uredinaceae. It is a common disease which occurs during the winter season (November-February). Developed leaves seem to be more susceptible to this illness.

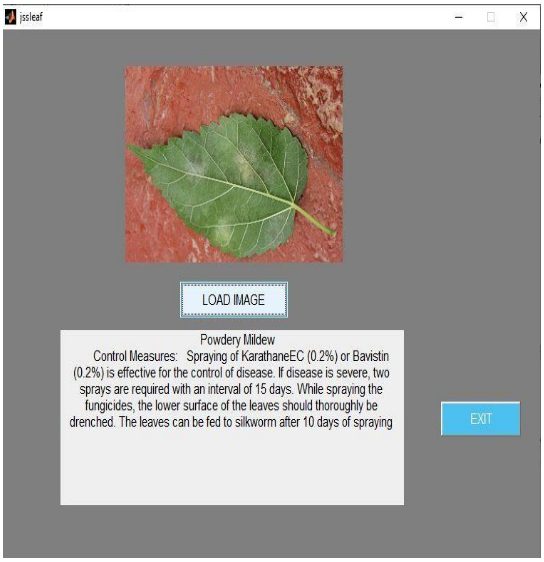


Fig. 12: Powdery Mildew

Figure 12 represents Powdery mildew. Phyllactiniacorylea karst tends to cause powdery mildew. This has been the most infectious illness and it spreads rapidly. The fungus corresponds to the Erysiphaceae group, category Ascomycetes Erysiphales order.

VII. RESULT AND ANALYSIS

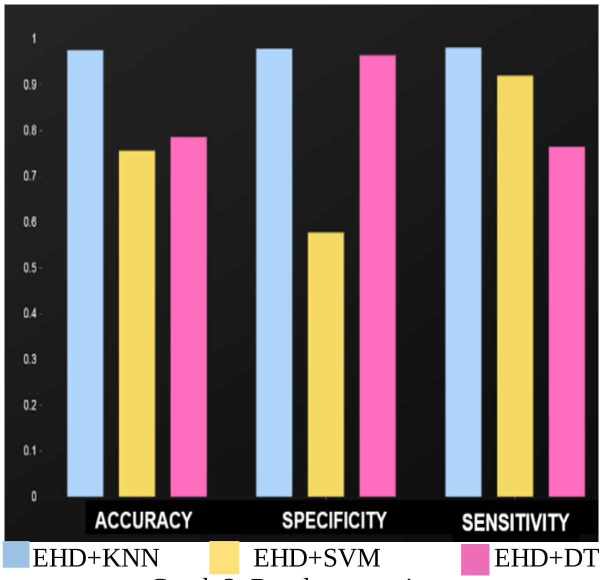
The disease classification performance analysis of the KNN classifier on leaf dataset. The leaf disease detection and classification result has been simulated on MATLAB. The segmentation output of the KNN classifier has been compared with the manually segmented result of the leaf image.

- **Accuracy:** Accuracy indicates the number of correctly predicted data points out of all the data points.
- ***Accuracy = (TP+TN) / (TP+FP+FN+TN)***
- **Specificity:** Specificity is the ratio between how much is correctly classified as negative to how much is the actual negative.
- ***Specificity = TN / (TN+FP)***
- **Sensitivity:** Sensitivity is a measure of how well a machine learning model can detect positive instances.
- ***Sensitivity = TP/(TP+FN)***

ALGORITHM	ACCURACY	SPECIFICITY	SENSITIVITY
EHD + KNN	97.50%	0.9778	0.9801
EHD + SVM	75.56%	0.5769	0.9192
EHD + DT	78.50%	0.9640	0.7645

Table 2: Result Comparison

ISSN No:-2456-2165



Graph 2: Result comparison

From the above result it is observed that EHD with KNN gives 97.50% accuracy, 0.9778 specificity. It is observed that for this combination of algorithm we have obtained better precision which indicates that the ratio of the correctly positively predicted data by our program to all positive data is accurate.

EHD with SVM gives 75.56% accuracy, 0.5769 specificity. It is observed that for this combination of algorithm we have obtained better precision which indicates that the ratio of the correctly positively predicted data by our program to all positive data is accurate.

EHD with DT gives 78.50% accuracy, 0.9640 specificity. It is observed that for this combination of algorithm we have obtained better specificity which indicates the true negative rate which means that the proportion of actual negatives, which got predicted as the negative.

VIII. CONCLUSIONS

In this paper, we briefly explained the Mulberry leaf disease detection. We were able to detect the disease using SVM and KNN algorithm. After the dataset is required and given as input, preprocessing takes place and will detect the mulberry leaf disease and their solution automatically once the disease is identified.Our experimental results indicate that EHD with KNN gives the best result with 97.50% of accuracy, 0.9778 of specificity.

ACKNOWLEDGEMENT

We are thankful to the resourceful guidance, timely assistance and graceful gesture of our guide Mrs. Nagaveni B Nimbal, Assistant Professor Department of Computer Science Engineering who has helped us in every aspect of our project work.

REFERENCES

[1.] 24th International Conference on Information Technology (IT) Zabljak, 18 – 22 February 2020 Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms DraškoRadovanović 1, Slobodan Đukanović 1, 2, Member, IEEE.

[2.] International Conference on Communication and Signal Processing, April 4-6, 2019, India Machine Learning for Plant Leaf Disease Detection and Classification – A Review L. Sherly Puspha Annabel, Member, IEEE, T. Annapoorani and P. Deepalakshmi.

[3.] 2018 International Conference on Computer, Control, Informatics and its Applications, Machine Learning-based for Automatic Detection of Corn-Plant Diseases Using Image Processing BudiariantoSuryoKusumo, Ana Heryana, Oka Mahendra, and Hilman F. Pardede, IEEE

[4.] 2019 International Conference 7-9 February, 2019, A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier Eftekhar Hossain1, Md. Farhad Hossain2 and Mohammad Anisur Rahaman3, IEEE

[5.] Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018), Diseases Detection of Cotton Leaf Spot using Image Processing and SVM Classifier Namrata R.