International Journal of Research in Engineering, Science and Management
Volume-2, Issue-5, May-2019
www.ijresm.com | ISSN (Online): 2581-5792

Segregation of Mulberry Leaves using Image Processing

M. Anusha[1], E. Harini[2], Shriram Subhas Hegde[3], Varrsha R. Kumar[4]
[1,2,3,4]Student, Dept. of Electronics and Communication Engg., The National Inst. of Engineering, Mysuru, India

***Abstract*: Silkworms mainly feed on mulberry leaves. Pests like fungus, virus and bacteria causes infection to leaves. Feeding diseased leaves to worms results in lower quality of the silk produced. So, there is a need to segregate diseased and healthy leaves, and feed only good and healthy leaves. This paper is designed to segregate these leaves using an intelligent monitoring platform framework. Automating the system reduces the need of labor, cost and efforts. Hardware and software integration makes the work easier. The major steps comprises of Image acquisition, image pre-processing, segmentation, feature extraction and classification. Color sensor is used to detect the leaf. Leaf being carried by the conveyer belt whose image will be captured by Raspberry-Pi camera module and processed. The captured image is pre-processed to remove noise. Combinations of color, texture and shape features are utilized to train the system, and the feature vectors are passed to K-Nearest Neighbor(KNN), Support Vector Machine(SVM) and Naive Bayes classifier algorithms for classification. Accuracy values are checked and tabulated for all the classifying algorithms. The algorithm with maximum accuracy, whose code is dumped into the Raspberry Pi for classification purpose. Servo motor being connected with the Raspberry Pi, does the segregation by rotating the plate attached to it, based on the class specified by the Raspberry Pi. KNN algorithm outperforms Naive Bayes and SVM. KNN gives the highest accuracy of 83.33% and hence KNN is considered for segregation.**

***Keywords*: Raspberry-Pi, healthy leaf, pre-processing, segmentation, feature extraction, KNN, SVM, Naives Bayes.**

1. Introduction

India is the second largest country to produce silk in the world. Hence, there is a vast scope for sericulture in India. Sericulture is an agro-based industry which involves cultivation of mulberry plants and rearing of silkworms. Silk is produced from the cocoons of the silkworms. The only food of silkworm is mulberry leaves. Pests cause infections to the plants. Hence before feeding the worms with leaves, the segregation of healthy and diseased leaves is necessary. Because, feeding diseased mulberry leaves results in bad cocoon formation, and hence the quality of the silk produced will be affected. Hence utmost care has to be taken. Seri culturists manually segregate the leaves before feeding silkworms. If the quantity of leaves is less, then there is no harm. But for the bulk quantity of leaves, manual segregation is waste of time, cost and labor. So it is necessary to upgrade the skills of the sericulturists to use the full potentialities, to produce qualitatively superior cocoons and to earn profitable income.

Fig. 1(a)

Fig. 1(b)
Fig. 1. Healthy and diseased leaves

This work proposes automating the process using Image processing. The major steps involved are image acquisition, pre-processing, segmentation, feature extraction and classification. The major task is to collect the Datasets of healthy leaves and diseased leaves. Figure 1(a) and Figure 1(b) shows few images of healthy and diseased leaves. These are used as training datasets by the system. The incoming leaves that to be segregated are the test datasets. Leaves are carried by the conveyor belt. Raspberry-pi camera module captures the image of the leaf (image acquisition) and sends it for the pre-processing check. Pre-processing involves resizing of the image, leaf detection and noise removal. Segmentation comprises of cropping the image to the specified dimension, thresholding and masking. Global descriptors like color, shape and texture are used for the feature extraction, where these vectors are extracted from both training and test images separately. These vectors are fed as input to the classifiers for the classification. KNN, SVM and Naive Bayes algorithms compute separately, comparing the training vectors and test vectors, and report the results. The accuracy values of all three classifiers are tabulated. The best classifier algorithm is loaded onto the Raspberry Pi for segregation.

2. Literature survey

Dataset collection is the main criteria for any image classification process. We visited Central Sericultural Research and Training Institute(CSRTI), and gathered information about the life cycles of silkworm, cultivation of mulberry leaves and other processes. We collected sufficient images of the mulberry leaves for training the system.

G. Sentilkumar et al. [1] has proposed an image capturing technique using embedded systems using Raspberry Pi. The whole system is composed of an image capturing camera, Raspberry Pi board to run programs on it. DVI compatible monitor is also connected to the system to display the captured image. The main processing unit connects the camera module and a display unit. The design rely on the SD card for long term storage and booting. The camera module used here is RPI NOIR camera board, which can deliver clear 5MP resolution images. This system can be operated in two sessions: firstly to capture the image and to create a database, and secondly to capture the image which can be used for identification and comparison. This system is smaller, lighter and consumes low power, thereby making it an effective system design for capturing the image. In the system proposed in this paper, whenever the leaf is detected, the image is captured, labelled, and saved in separate folder, and this image is fetched for further processing.

Jayaprakash Sethupathy and Veni S., [2] proposed a Disease identification method in mango leaves for Indian agriculture, using OpenCV. OpenCV implementation provides easily usable applications for the farmers. The system proposed in this paper, uses OpenCV machine learning software library for software implementation. It has more than 2500 optimized algorithms, which are used to detect objects, identify faces etc. It has Python, C++, Java and MATLAB interfaces.

Nikhil Govinda Dhakad et al. [3] has proposed an automatic disease detection of plant leaf using image processing techniques. It involves image acquisition, image pre-processing, image segmentation, feature extraction and classification. This work says that Image processing is the advanced computer technology, using which an automatic detection system is developed, supports farmers to identify the diseases at an initial stage. The Controller successfully monitors the humidity and temperature, and stores the sensor parameters in the timely manner. This will help the users to analyze the various parameter conditions in the field, anytime anywhere, which makes an automatic disease detection system is more efficient the traditional detection process.

In the system proposed in this paper, the image captured is passed through all the stages mentioned in the above paper for proper detection of abnormalities in leaves, and finally to output about the status of leaf (diseased or healthy).

Arti N. Rathod et al. [4] suggested in agriculture research of automatic leaf disease detection is a benefited topic in monitoring crops in large fields, thereby detecting the symptoms of diseases in plant leaves automatically, in the early stages. This proposed work initially captures the image, remove noises using median filter and convert the RGB image into CIELAB color component. Secondly, the image is segmented using k-medoid technique. Thirdly, masking green-pixels and removal of masked green pixels, and calculating the texture feature statistics. Finally, this features are passed into neural network. The neural network classification successfully detect and classify the tested disease.

In the system proposed in this paper, concept of masking, color space conversion and texture feature extraction are used as a part of segmentation.

Sachin B. Jagpat and Shilesh M. Hambarde [5] proposed crop diseases are indicated by leaf spots. These spots are examined and subjected to knowledgeable opinion. This system consists of four stages; the first is image enhancement, which includes HIS transformation, Analysis of histogram, and intensity adjustment. Segmentation would be the second stage. Feature extraction being the third stage, includes three feature descriptors namely color, size and shape. The fourth stage is classification, which uses back propagation based neural network system. In the system proposed in this paper, an integrated image processing system is developed to detect the spots and abnormalities on leaves, to identify whether the leaf is affected by disease or not, as a part of segmentation.

T. Thamil Azhagi et al. [6] proposed a methodology where the leaf image is captured, pre-processed, segmented and classified. Image pre-processing involves masking, where image pixel value is set to zero or any other background value. The pixels with most green colored are identified here. Specific threshold value will be computed for the pixels. The pixel components red, green and blue are giving zero value by RGB component mapping.

P. B. Chopade and Katkar Bhagyashri P., [7] proposed a system for identifying and classifying the leaf diseases affecting the fruit crops which uses image processing tool. All the information regarding the disease is sent to Farmer’s phone through GPRS. This work mainly focuses on image segmentation, where the intensity difference between infected leaf concepts is used to segment. Pre-processing stage’s output is fed to segmentation, where segmentation is carried out based on thresholding and morphological operations.

In the system proposed in this paper, lower and upper threshold value is calculated from histogram image of saturation image and then the saturation image is divided into two binary images based on threshold values.

International Journal of Research in Engineering, Science and Management
Volume-2, Issue-5, May-2019
www.ijresm.com | ISSN (Online): 2581-5792

Aniket Gharat et al. [8] proposed a leaf disease detection model, which makes of image processing and convolutional neural network(CNN). Image feature extraction is the main criteria here. Features like color(HSI model) and shape(Hu moments) are used. Classification based on disease is done by CNN. The system uses two image databases, one for training images and the other for test images.

In the system proposed in this paper, three global descriptors like color, shape and texture, with two datasets(train and test) are considered. Color uses histogram and shape uses Hu moments, for feature extraction. These features are extracted and then given to the classifier for classification.

Jinko Kim et al. [9] proposed two frameworks for image classification. KNN being the simplest learning algorithm, classify the objects based on the closet training examples in feature space. Training process in this algorithm involves labelling the training images and storing the featured vectors. The test image(unlabeled image) is assigned to the label of its k nearest neighbours, where based on the k value specified the image is classified. SVM classifier is the supervised machine learning algorithm uses a technique called kernel. Based on the experiment, SVM outperforms KNN.

In the system proposed in this paper, three classifiers two of which KNN and SVM are used for the classification purpose. These two algorithms provides better classification even though the training dataset is less.

Khumukcham Robindro et al. [10] proposed a methodology to diagnose the disease in rice plant which uses a concept of decision Support System(DSS) using Naive Bayes. It contain necessary rules on predicting decisions on diseases, which also offers solution how to handle the predicted disease. This paper provides necessary details on Naïve Bayes analysis. In the system proposed in this paper, KNN, SVM and Naïve Bayes are used as classifiers and efficiencies of these three classifiers are compared.

3. Architecture

The smart agriculture model aims to detect healthy and diseased leaves. It is of low cost and has good efficiency. Figure 2 shows the architectural design of the proposed system. It includes PC, raspberry-pi, camera module, relay module, DC motor and servomotor.

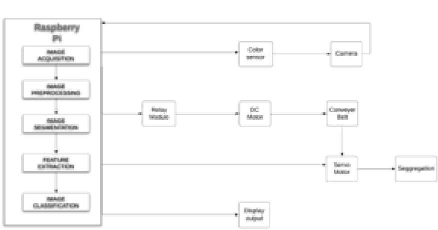

Fig. 2. Architecture

Raspberry Pi is connected to color sensor, relay module and camera module. The leaf is being carried by the conveyor belt. Whenever the leaf is detected by the sensor, image is captured by the camera module [1]. The captured image is stored in a separate folder and sent for further processing. Once the processing is done, servo motor rotates as programmed and segregates the leaf.

A. *Methodology*

The system mentioned in this paper is the combination of both hardware and software implementation. Hardware implementation is mentioned in the architecture section. In the software section, required datasets are collected and stored in the separate folder called Training folder. This folder contains both healthy leaves and diseased leaves. These datasets are also called as training datasets, which are used to train the system, whose morphological characteristics and features are extracted and compared with the test images to obtain the result. The major steps of the image processing are shown in Figure 3.

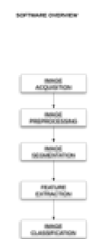
Fig. 3. Steps of the image processing

B. *Image acquisition*

Image acquisition is the process of capturing the image of the input data. In the proposed system, leaves are carried by conveyor belt. Color sensor is used to detect the leaf. Whenever leaf is detected, motor stops the movement of conveyor belt and the camera module captures the image[1]. This unlabeled image will be stored in a separate folder called Test folder. These datasets are called as test datasets. The system is designed in such a way that, any number of images can be captured, until the camera is released. The captured RGB image is converted into HSV color space providing the upper and lower limits to detect green pixels in the image. In the absence of green pixels, the image will not be captured. Figure 4 shows the captured image of the unlabeled leaf.


Fig. 4. Unlabeled leaf

C. *Image pre-processing*

Image pre-processing is the method of improving the quality of the image and removing the noises in the image. The captured image should be processed before sending through other steps. The input images will be in different sizes and qualities. To compare the test image with the training datasets, both dimensions should match. Hence resizing of image is mandatory, otherwise it throws error. The test image is converted into gray scale image and reduced to the specified dimension, matching the dimensions of training datasets.

D. Image segmentation

Image segmentation is used to convert the digital image into different segments having a few similarities [5]. This helps in leaf detection and boundary line detection in the image. In the proposed system, segmentation involves cropping the image, masking the image and contouring the image.

1) Contouring

Contour represents the shape of the boundaries of the object in an image. The input to the contour finding process is a binary image which is obtained by applying thresholding [7]. The region of interest(region to be detected) in the binary image should be white and the background of the image should be black. In the proposed system, the region to be detected is the green pixels of the leaf. Figure 5 shows the circle contouring the leaf part detecting the green pixels.

Fig. 5. Circle contouring the leaf part detecting the green pixels

2) Masking

Masking is the process of setting some pixel values in an image to zero and others to non-zero [6]. The Image masking tool accepts the input image, masks it and produces a new image which is a copy of the input image, except that the new image will have its pixel intensity values set to zero according to the masking operation. This tool performs opening (eroding followed by dilating function) and closing(dilating followed by eroding function) using kernel function. The kernel function is a structuring element having elliptical kernel of size 5x5. Figure 6 shows the masked image.

Fig. 6. Masked image

4. Cropping

Cropping the image retains only the region of interest where, everything else will be removed. In the proposed system, base and tip of the leaf are considered as the extreme points in vertical direction. Similarly extreme points in the horizontal direction are considered. This is done using OpenCV [2] inbuilt tools like OTSU and THRESH. This process makes sure that redundant information in the image are excluded. Figure 7 shows the cropped image.

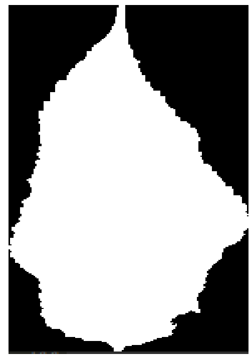

Fig. 7. Cropped image

A. Feature extraction

Features are the information that is extracted from an image. Features are the real valued vector numbers which describe the image at high level perspective. In the proposed system, Global feature descriptors like color, texture and shape are considered for the detection. The color descriptor used is Color Histogram [8]. Histogram is a graphical representation of number of pixels in an image. It quantifies the pixel value for each intensity value. It yields 8 vector values(bin=8).

The texture descriptor used is Haralick [4]. It yields 14 vectored values, out of which only 13 vectors are considered, since the last vector value need high computation time. The shape descriptor used is HuMoments [8]. It is invariant to various image transformations. It yields 7 vectored values. All these features are calculated for both training and test datasets, and compared to obtained results.

B. Classification

The feature vectors calculated in the previous steps, become the input to the classifiers. Classifiers like KNN, SVM and Naïve Bayes are used for the computation. They compare the test featured vectors with each training featured vectors. The difference is calculated with respect to healthy and diseased image featured vectors. The test image is classified to the class (healthy or disease) which has least difference.

KNN is one of the simplest machine learning algorithms [9] opted for classification and regression. The value k is the deciding factor. K value is usually odd for determining between two classes.

SVM [9] locate decision boundary which divides data points in the space into two classes. There can be any number of hyperplanes, but the one with maximum distance from the data points is optimal. SVM considers the output in linear function. If the value is greater than 1, then it classifies as class1. If the value is -1, then it classifies as class2.

Naive Bayes algorithm has a probabilistic approach [10]. It calculates the probability for all the input vector values and selects the value with maximum probability value. This value is computed for conditional and class probability, and compared with the training vector values for the classification.

5. Experimental result

The servomotor motion decides to which class the input leaf belongs. If the leaf is healthy, servomotor rotates to the left side (90degree to 0degree). If the leaf is diseased, servomotor rotates to the right side (90degree to 180degree). After classification, the accuracies of all three algorithms are calculated. KNN exhibits more accuracy compared to the other two algorithms.

The collected datasets are divided into different sets of training and test datasets. The accuracies for all different combinations, and for all the three classifiers are computed and tabulated. Figure 8 shows the accuracy tabulation for different dataset combinations.

Algorithm	Train/Test Size	No. of Images	Accuracy
KNN	80/20	24:6	83.33333333333334
	75/25	22:7	75.0
	70/30	21:9	66.66666666666666
	50/50	15:15	86.66666666666667
	40/60	12:18	94.44444444444444
SVM	80/20	24:6	33.33333333333333
	75/25	22:7	87.5
	70/30	21:9	33.33333333333333
	50/50	15:15	40.0
	40/60	12:18	88.88888888888889
NAIVE'S BAYES	80/20	24:6	66.66666666666666
	75/25	22:7	62.5
	70/30	21:9	77.77777777777779
	50/50	15:15	80.0
	40/60	12:18	77.77777777777779

Fig. 8. Accuracy tabulation for different dataset combinations

6. Conclusions

Detecting whether the leaf is healthy or diseased is important. The proposed methodology provides an automated system for segregation of leaves which reduces cost, time and labor. This has been implemented successfully using OpenCV Python. OpenCV implementation provides easy availability for the usage of farmers in the form of low cost application specific system. Hardware and software combination makes the system simple and efficient. Leaves are carried by conveyor belt, whose images are captured by raspberry-pi camera module, and sent for image processing. This system uses KNN, SVM and Naive Bayes classifying algorithm. Based on several combinations of color, texture and shape features, classification is performed using above mentioned classifiers and accuracies are compared. The maximum accuracy value obtained is 83.33%, for KNN classifier. KNN works better even if the training datasets are less, where other algorithms need huge amount of data clusters. Servomotor direction indicates the result.

References

[1] G. Senthilkumar, K. Gopalakrishnan and V. Sathish Kumar, "Embedded image capturing system using Raspberry Pi", International journal of emerging trends and technology in computer science (IJETTCS), Volume 3, Issue 2, March – April 2014.

[2] Jayaprakash Sethupathy and Veni S., "OpenCV based disease identification of mango leaves", Intenational Journal of Engineering and Technology(IJET), Volume 8, No 5, Oct-Nov 2016.

[3] Nikhil Govinda Dhakad, Umesh Shyamkant Yewale, Tejal Yuvaraj Patil and Gayatri Avinash Deore, "Leaf disease detection using image processing", International Research Journal of Engineering and Technology(IRJET), Volume 5, Issue 12, December 2018.

[4] Arti N. Rathod, Bhavesh A. Tanawala and Vatsal H. Shah, "Leaf disease detection using image processing", International Journal of Advance Engineering and research Development(IJAERD), Volume 1, Issue 6, June 2014.

[5] Sachine B. Jagtap and Shailesh M. Hambarde, "Agricultureal plant leaf disease detection and diagnosis using image processing based on morphological feature extraction", IOSR Journal of VLSI and Signal Processing(IOSRJVSP), Volume 4, Issue 5, Sept.-Oct. 2014.

[6] T. Thamil Azhagi, K. Swetha, M. Shravani and A. T. Madhavi, "Plant pathology detection and control using Raspberry Pi", International Journal of Engineering Sciences & Research Technology, March 2018.

[7] P. B. Chopade and Katkar Bhagyashri P., "Image processing based detection and classification of leaf disease on fruit crops", Special issue on International Journal of Electrical, Electronics and Computer Systems, For 3rd National Conference on Advancements in Communication, Computing and Electronics Technology(ACCET-2016).

[8] Aniket Gharat, Krupa Bhatt, Bhavesh Kanase and Abhilasha Bapnna, "Leaf disease detection using image processing", Imperial Journal of Interdisciplinary Research(IJIR), Volume 3, Issue 5, 2017.

[9] Jinko Kim, Byung-Soo Kim and Silvio Savarese, "Comparing image classification methods: k-nearest-neighbor and support-vector-machine", Applied Mathematics in Electrical and Computer Engineering.

[10] Khumukcham Robindro, Ksh. Nilakanta, Ashoke Das and Deepen Naorem, "Decision Support System for Rice Plant Disease Diagnosis using Naïve Bayes' Algorithm", International Journal of Scientific Research in Multidisciplinary Studies, Volume 3, Issue 7, July 2017.