Mango Leaf Diseases Recognition Using Neural Network and Support Vector Machine

Md. Rasel Mia
Department of Computer Science and Engineering
Gono Bishwabidyalay
Dhaka, Bangladesh
rasel376mahmud@gmail.com

Sujit Roy
Department of Computer Science and Engineering
JKKNIU and Gono Bishwabidyalay
Dhaka, Bangladesh
roysojib09102029@gmail.com

Subrata Kumar Das*
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
sdas_ce@yahoo.com

Md. Atikur Rahman
Department of Computer Science and Engineering
Gono Bishwabidyalay
Dhaka, Bangladesh
atikhasan.cse@gmail.com

Abstract—*This paper presents a Neural Network Ensemble (NNE) for Mango Leaf Diseases Recognition (MLDR). Mango trees are affected by various diseases and identifying disease is a complex task till now because those diseases are detected manually. This study intends to detect the symptoms of plant diseases easily with machine learning than a manual monitoring system. Here, Trained data are produced by classification technique collecting images of leaves that were various disease affected. A machine learning system is designed to identify the symptom of mangoes leaf diseases automatically uploading and matching new images of affected leaf with trained data.The proposed system could successfully detect and classify the examined disease with average accuracy of 80%. This proposed solution would clinch the Mango plants. The system will help to detect disease without the presence of agriculturist. It would also save time to identify disease with machine instead of traditional system, which helps to treat the affection of mango leaf disease properly, increase the production of mango and meet the demand of global market.*

Keywords-Artificial Neural Network (ANN); Disease; K-Nearest Neighbor (KNN); Mango plant; Pattern recognition; Support Vector Machine (SVM)

I. INTRODUCTION

Mango plants are very important to keep biodiversity and it is a major source of fruits. That's why Mango plants cultivation has been a growing awareness over the world for promoting fruit production in a sustainable agricultural practices. But Mango plants disease is a great barrier to produce enough fruits to meet the people demand. There are many diseases, like Mango malformation disease, Anthracnose, Bacterial flower disease, Golmachi, Moricha disease, Shutimold, Bacterial black spot, Apical bud necrosis, Red rust, Lichens, Powdery mildew, Root rot and damping off, Ganoderma root rot, which affect the mango trees. Numerous Mango trees of majority countries over the world are suffered by Powdery mildew and damaged upto 23% unsprayed trees[1][2][3]. Losses of mango trees are caused upto 39% worldwide by Anthranose[4][5]. Kumar and Beniwal cited that Mango malformation affects over 50% trees in india including Egypt, Africa, Mexico, America, Sudan, Cuba, Australia, Bangladesh [6][7]. Mango Bacterial Canker is a fatal disease which causes 10-100% loss to mango yield in the field as well as in storage[8][9]. Smith reported that around 93% mango plants in some orchards are suffered from Bark Cracking[10]. So, Mango diseases cause a crucial problem and results in economic and agricultural industry loss. If those diseases are possible to identify rightly then the prevention of disease would be easy.But, the identification of diseases is so difficult with human naked eye. Machine learning could be a feasible detection way of MLD for the advancement in technology, especially for the development of image processing technique.

Pattern Recognition is an affiliate of machine learning that focuses on the recognition of pattern and discipline in data. Pattern Recognition system have supervised and unsupervised learning data.It is closely related to artificial intelligence and machine learning, organized with applications such as data mining, fruit/vegetable recognition, diagnostic systems, face recognition, biometrics, image processing, classifying galaxies by shape, identifying fingerprints, highlighting potential tumors on mammogram, handwriting recognition and diseases recognition etc. Techniques from statistical pattern recognition have obtained a widespread use in digital image processing since the revival of neural networks.

The evaluation and development of leaf diseases has become more common nowadays, as environment and climate condition are more unstable than ever. In this changing environment, appropriate and timely disease identification including early prevention has been more important.

The organization of this document is as follows. Related work is discussed in section 1, Methodology and Design of the MLDR in section 2, Result and Discussion in section 3, and Conclusion and Future Work in section 4.

II. RELATED WORK

The infection of diseases in plant or tree is a great problem in agricultural sectors. As a result, the leaf or bud diseases recognition is a popular area in recent years. The classification of plant leaves is a crucial process in botany and in tea, cotton and other industries.

Warne and Ganorkar describes the approach to prevent the crops from heavy loss by careful detection of disease[11]. Shergill et al. elaborate an approach is useful in crop protection especially in large area farms, which is based on automated techniques that can detect diseased leaves using color information of leaves[12].

Ranjan et al. [13] describes a diagnosis process that is mostly visual and requires precise judgment and also scientific methods. Image of unaffected and diseased leaf is captured. As the result of segmentation Color HSV features are extracted, artificial neural network (ANN) is then trained to distinguish the healthy and diseased samples. ANN classification performance is 80% better in accuracy. Kajale describes the approach for detection and computation of texture information for plant leaf diseases [14]. The processing system consists of four main steps, color image is converted to HSI, then the green pixels are masked and removed using specific threshold value, then the pre-processed image is segmented and the useful segments are extracted, finally the texture information is obtained. The diseases present on the plant leaf are evaluated based on the texture information.

Al-Bashish et al. and Ghaiwat et al. [15][16] presents investigation on different classification methods that can be used for plant leaf disease classification. For given investigation pattern, k-nearest-neighbour method is seems to be suitable as well as simplest of all algorithms for class prediction. If training data is not linearly separable then it is difficult to determine optimal parameters in SVM, which performs as one of its drawbacks.

Garg et al.presents analysis on a support vector machine is a very potential AI method and can apply extensively to solve classification problems [17]. The SVM that is used to solve regression problems is known as support vector regression (SVR). SVR is very popular among researchers for providing generalization ability to the solution model. Ullagaddi and Raju explained their idea in a conference to detect disease of Mango crop using modified rotational kernel transform features[18]. Bhange & Hingoliwala, and Kaur et al. worked on disease detection and classification for showing how to lead smart farming [19][20]. In 2017, Madiwalar & Wyawahare conducted a comparative study on plants disease identification to explain the overall view of our recent and previous works on it in a conference [21]. Dai et al. and Rump et al. did work taking classification algorithm, and initial plant disease classification & detection [22][23]. TIhobogang's and Wannous's recently displayed their ideas in a conference paper for designing of plant disease detection system [24]. In 2018, Kulkarni et al. [25] presents a methodology consists of color transformation, masking of green pixels using thresholding, segmentation, texture feature extraction using gray level co-occurrence matrix GLCM and classification.

In this paper, we present a Neural Network Ensemble (NNE) for detecting mango leaf disease. Apart from this, four types of diseases are recognised pragmatically and tested their accuracy.

III. METHODOLOGY AND DESIGN OF THE MLDR

In order to develop MLDR among various NNE techniques, SVM has been employed. The SVM technique along with efficient feature extraction enables MLDR to attain faster training and recognition performance.

A. Proposed Model

When a new image inserted into the system that new data would be compared with trained data. The workflow of proposed MLDR architecture is shown in “Fig. 1”.

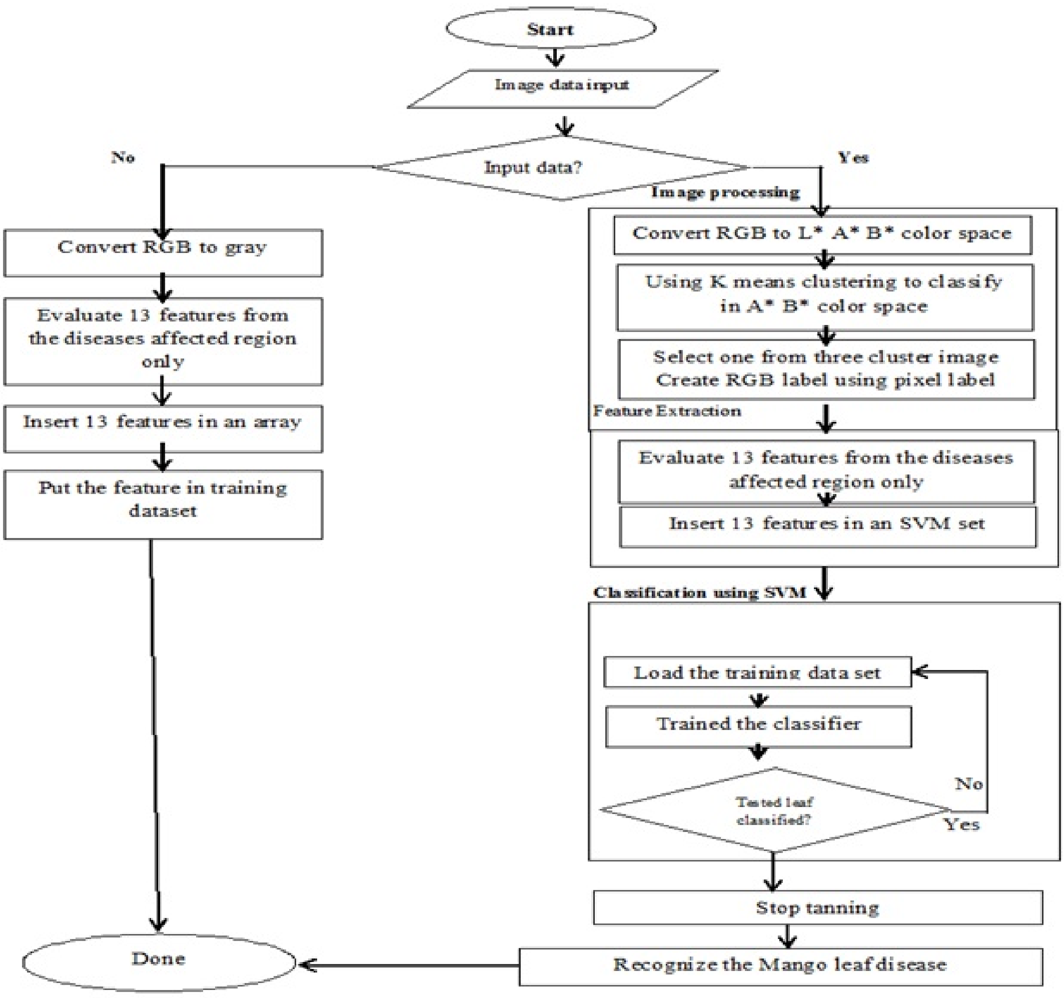

Figure 1: Analytical approach of MLDR

B. Mathematical Analysis

Gray-Level Co-occurrence Matrix (GLCM) is a-level spatial dependence matrix. It is frequently used in the literature without a hyphen, co-occurrence. Gray Level Co-occurrence Matrix creates the GLCM =Graycomatrix(I). Here gray scale intensity value I occurs horizontally adjacent to a pixel with the value j. Each element (I, J) in GLCMS specifies the number of times that the pixel with value I occurred horizontally adjacent to a pixel i^{th} value j. We will evaluate some features like Contrast, Correlation, Energy, and Homogeneity by using GLCMS.

Contrast: Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = [0 (size (GLCM, 1)-1) ^2]. Contrast is zero for a constant image.

Formula: $\sum_{i,j} |i-j|^2 p(i,j)$

Correlation: Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] Formula: $\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$

Correlation is one or -one for a perfectly positively or negatively correlated image. Correlation is Null for a constant image.

Energy: Returns the sum of squared elements in the GLCM Range = [0 1]. Energy is one for a constant image. Formula: $\sum_{i,j} p(i,j)$

The property Energy is also known as *uniformity*, *uniformity of energy*, and *angular second moment.*

Mean: M = mean (A) returns the mean of the elements of A along the first array dimension whose size does not equal 1. Formula: $X = \frac{\sum_{i=1}^{n} x_i}{n}$

Standard Deviation: For a random variable vector, A made up of N scalar observations. The standard deviation is the square root of the variance. Some definitions of standard deviation use a normalization factor of N instead of N-1, which you can specify by setting w to 1. Formula: $\sigma = \frac{\sum(x_i)}{n}$

C. *Different Types of Diseases*

There are many diseases that can affect the mango leaf mentioned above. If disease infects the leaf, stems and fruits, the leaf become yellow and then brown is shown in “Fig 2”. After the corner of the leaf, become dry. One kind of insects attracts its new leaf and for this reason, there have some gol in the new leaf.

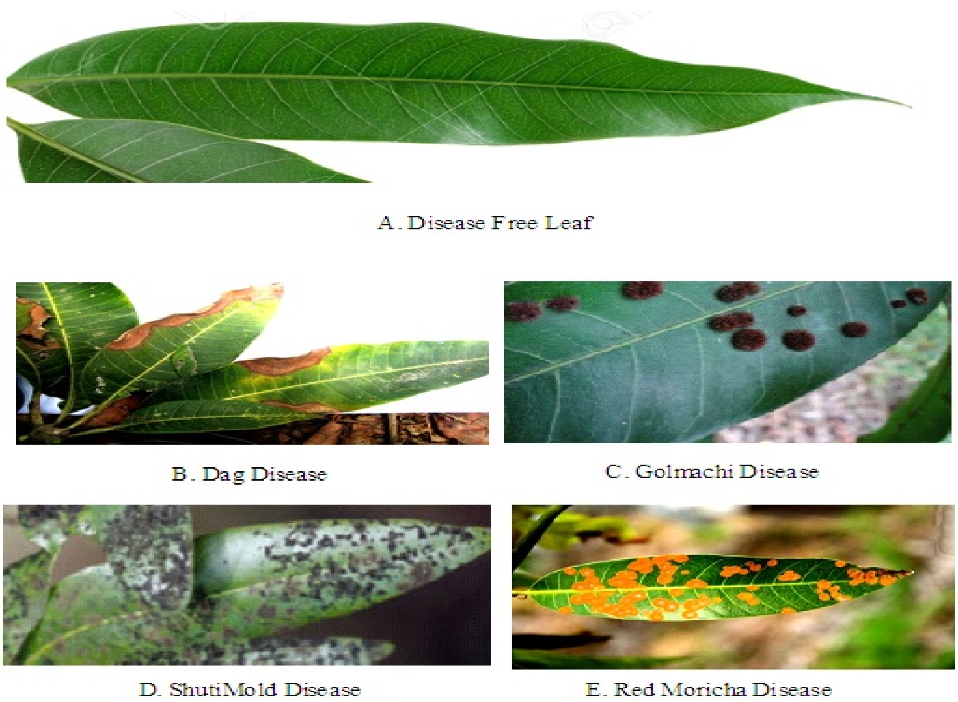

Figure 2: (A) Disease Free Leaf, (B) Dag Disease, (C) Golmachi Disease, (D) ShutiMold Disease, and (E) Red Moricha Disease

The attack on the disease, resulting in the trunk of the red and rust spots can be seen as a kind of high. A kind of green algae is to attack the disease. Most of time Mango’s shutimold disease can be seen in the presence of ant.

If the mangos hopar or milibag attack more, the invasion of mold fungi shuti occurs, the accumulation of dirt on the period and the branch dies and leaves the tip of the storm.

D. Input Data (Image) Analysis

The inputted data (image) are needed to classify to recognize the disease. For the recognition we have used SVM classification.

Image pre-processing: Firstly, we have taken the image from current directory. For getting better performance we can edit the image using any kind of photo editor (just like Adobe Photoshop). Actually the photo editor work is to increase or decrease of contrast of brightness or resizing of the image or rotating the image.

Classify the Color in A* B*: We have to convert the image from RGB to lab. For this we have to use srgb2lab () function. This function performs the work. Here it creates three images. Those images are being resized according to their previous size. It makes three color images. One is less brightness. Another is according to A*(indicating where the color falls along red-green) is presented in “Fig. 3”. And final is according to B*(indicating where the color falls along the blue-yellow).

Figure 3: Steps of A* B* image processing

Feature Extraction: Feature plays a very important role in the area of image processing. Before getting features, various image pre-processing techniques like contrast, correlation, energy etc. homogeneity are applied on the sampled image.

For SVM classification firstly we had to load .mat type file, which contained data sets. Then we had to run svmtrain(train, group), train consisted of our training data(Fig. 4). Group consisted of normal, dag disease, golmachi, moricha disease, shutimold. Then finally we had to run a method for the classification named svmclassify(svmstruct,sample).

diseasefeat

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.2117	0.8942	0.2283	0.8970	97.8807	39.7072	7.2376	15.8643	1.2097e+03	1.0000	2.3831	-0.3878	255
2	2.0000e-04	9.0000e-04	2.0000e-04	9.0000e-04	0.0979	0.0397	0.0072	0.0159	1.2097	1.0000e-03	0.0024	-4.0000e-04	0.2550
3	2.0000e-04	1.0000e-03	3.0000e-04	9.0000e-04	0.0788	0.0557	0.0072	0.0159	2.6236	1.0000e-03	0.0055	0.0017	0.2550
4	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1244	0.0957	0.0064	0.0148	6.3283	1.0000e-03	0.0015	2.0000e-04	0.2550
5	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.0929	0.0755	0.0076	0.0155	5.3120	1.0000e-03	0.0019	6.0000e-04	0.2550
6	1.0000e-04	9.0000e-04	2.0000e-04	9.0000e-04	0.0929	0.0435	0.0074	0.0159	1.2368	1.0000e-03	0.0026	4.0000e-04	0.2550
7	1.0000e-04	9.0000e-04	2.0000e-04	9.0000e-04	0.1011	0.0477	0.0477	0.0159	2.0199	1.0000e-03	0.0025	1.0000e-04	0.2550
8	2.0000e-04	9.0000e-04	2.0000e-04	9.0000e-04	0.1306	0.0403	0.0073	0.0160	1.1756	1.0000e-03	0.0029	-3.0000e-04	0.2550
9	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1449	0.0699	0.0078	0.0157	4.0857	1.0000e-03	0.0021	-2.0000e-04	0.2550
10	0.1727	0.8718	0.2398	0.9152	76.4812	36.7481	6.8051	15.9686	783.8659	1	2.3210	0.6905	255
11	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1029	0.0694	0.0076	0.0147	2.8452	1.0000e-03	0.0022	2.0000e-04	0.2550
12	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1431	0.0745	0.0077	0.0156	4.5448	1.0000e-03	0.0020	-2.0000e-04	0.2550
13	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1449	0.0699	0.0078	0.0157	4.0857	1.0000e-03	0.0021	-2.0000e-04	0.2550
14	0.1727	0.8718	0.2398	0.9152	76.4812	36.7481	6.8051	15.9686	783.8659	1	2.3210	0.6905	255
15	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1029	0.0694	0.0076	0.0147	2.8452	1.0000e-03	0.0022	2.0000e-04	0.2550
16	2.0000e-04	1.0000e-03	1.0000e-04	9.0000e-04	0.1431	0.0745	0.0077	0.0156	4.5448	1.0000e-03	0.0020	-2.0000e-04	0.2550

Figure 4: Example of training data

IV. U RESULTS AND DISCUSSION

The classifiers are trained and tested using images of Mango leaf diseases. The sample images are divided into two halves, one-half is used for training, and other is used for testing. The color and shape features are used to train and test neural network model. To evaluate classification performance, our experimental study focuses on the comparison between our MLDR and SVMs that is trained by the single-step learning approach. As mentioned earlier, MLDR has been trained by 20 samples.

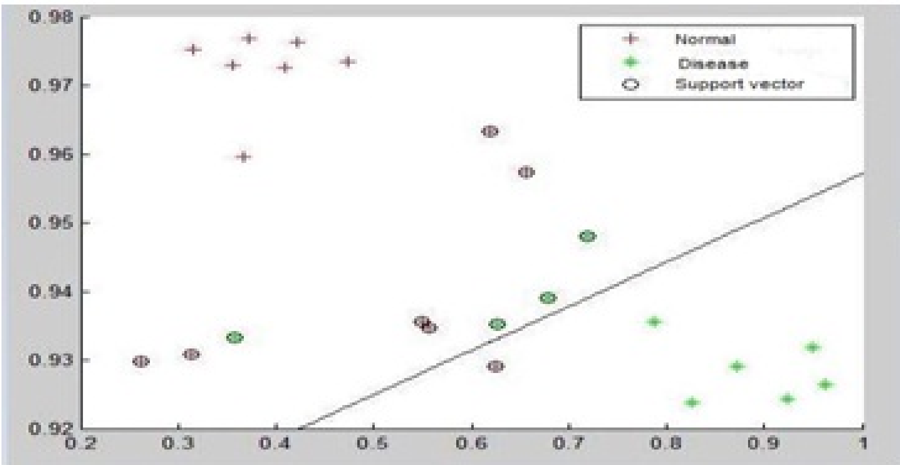

Figure 5: SVM classification between normal data and disease data

In “Fig. 5” there are two types of data, one is normal and other is disease. The straight line indicates the output data, which is very close to input data. The classifier chooses the approximate value for output. When a new image is selected as an input data, if the feature of new data is matching with the trainee data set whose are located in the middle of the graph is the approximate output data.

The free “+” in the plot indicates the data which are not near of input data of normal data’s class (support vector). And the adjacent of “+” and “o” are near of input data (support vector).

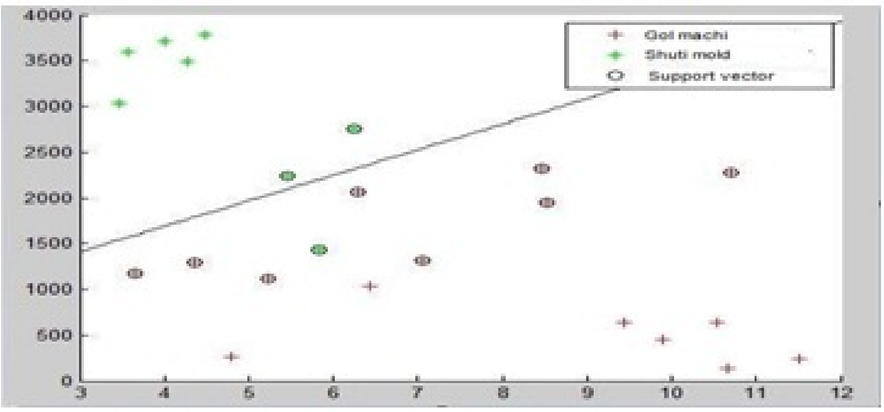

Figure 6: SVM Classification between Golmachi and Shutimold data

In “Fig. 6” there are two types of data, one is Golmachi and other is Shutimold. The straight line indicates the output data which are very close to input data.

The grayscale image and binary image of our input data are shown in “Fig. 7” and “Fig. 8” accordingly. In the image, the black spot only shows the disease affected area and rest of the image show the disease free area.

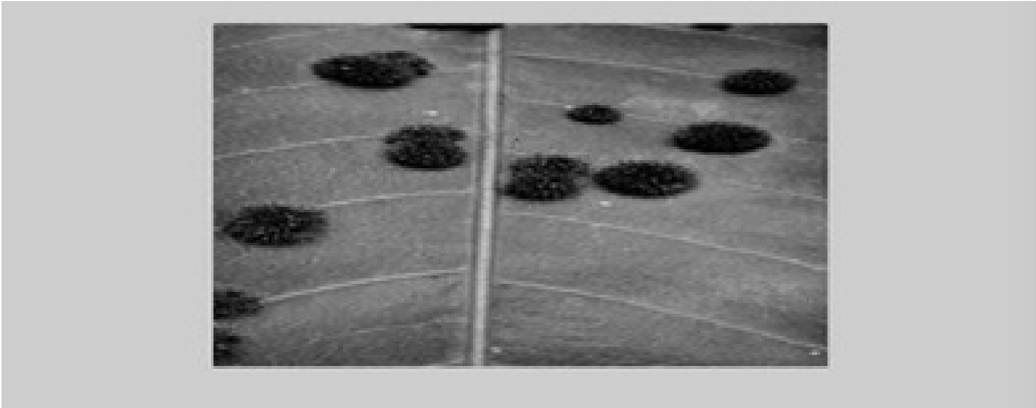

Figure 7: Grayscale image of the input data.

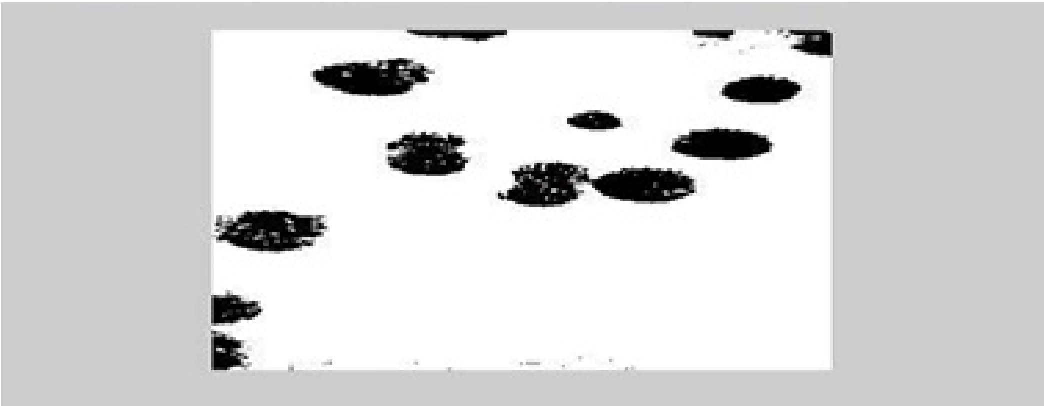

Figure 8: Binary image of the input data.

This system recognizes normal data, which are not diseases, affected with accuracy of 87.5%. We have taken 8 images as our input data in classification between normal and disease data. We got seven correct data and one incorrect data.

Accuracy = (total correct output data)/(total input data)*100%

So our accuracy is 7/(8)*100% = 87.5%.

By using support vector machines (SVM) we can check only two groups. Firstly, we have checked if the input data is disease free or not. The same process is used for disease affected leaf (Dag disease, Golmachi, Moricha disease, shutimold) and we have got different accuracy for using support vector machine (SVM) is shown in Table 01. The average accuracy is founded 80%.

TABLE I. PERFORMANCE OF MLDR

Input type	**Accuracy**	**Average accuracy**
Normal	87.5%	(87.5+75+76+75+86.5)/5=80%
Dag disease	75%	
Golmachi	76%	
Moricha disease	75%	
Shutimold	86.5%	

As mentioned earlier, MLDR has been trained by 20 samples. Here, training has been continued until the error rate has reached to 0.2. To gain the better performance, we have chosen 1000 training cycles for our experiment.

The accuracy of classification varies from 80%-82% depending on the algorithms and limitations of image acquisition. Classification also is obtained with great accuracy as the case with image detection. In this case, also the classification accuracy can be obtained up to 80% with correct imaging techniques and algorithms.

V. CONCLUSION AND FUTURE WORK

Bangladesh is an agriculture-based country. Its economy is solely dependent on agriculture. The food inflation is a national challenge. Scientists, agriculturists work day and night to promote the yield of food grains. It is very difficult to infer the varieties of a leaf disease by simple visual observation. It is very time consuming and can be accomplished by the trained botanists. Research outcome can help to recognition of Mango leaf diseases quickly and easily by machine instead of manual system.

After observing and classifying all these data, for getting more accuracy we need more data. For matching the disease we need various types of data just like left rotate, right rotate, up rotate, down rotate, 100% , 120%, 140%, 160%,180%, 200% zoom out.

In this primary stage, we have detected four types of disease and one normal among total five groups. Each group consists of four different types of images. For getting more accuracy, we need to add more images in each group. The accuracy depends on training data. In our work, we take 20 training data having 13 features each. To improve performance, we need to add more training data. Only for classification, we have used SVM classifier. We have to use the classifier more sufficiently for increasing efficiency.

Future work will include research along two directions: first comparing textures base features, second color features for improving recognition accuracies.

REFERENCES

[1] R. A. Peterson, G. L. Schipke, and C. P. Clarkson, "Significance of two mango flower diseases in the dry tropics," *In: Proc. Third Int. MangoSymp, held at Darwin*, pp. 338-343, 1991.

[2] Om. Prakash, K. A. Misra, and R. Kishun, "Some threatning diseases of mango and their management," In: Management of Threatning Diseases of National Importance. (Ed. V. P. Agnihotri, A. K. Sarbhoy and D. V. Singh) Indian Phytopatholigical Society.Malhotra Publication House, New Delhi, pp.179-205, 1997.

[3] A. V. Bourbos, and T. M. Skoudridakis, "First report of Oidium mangifereae on Mangifera indica in Greece," Plant Dis. vol. 79, pp. 1075, 1995.

[4] S. H. Sohi, , S. S. Sokhi, and P.R., "Gloeosporioides Penz. And its control," Phytopath. Mediterranea. vol. 1291, pp. 114-116, 1973.

[5] Om. Prakash, K. A. Misra, and K. B. Pandy, "Anthracnose disease of tropical and subtropical fruits," In: Disease Scenario in Crop Plants. VoU Fruits and Vegetables (Eds. V.P. Agnihotri, Om Prakash, Ram Kishun and A.K. Misra), Intern. Books and Periodicals Supply Service, Delhi, pp. 1-27, 1996.

[6] J. Kumar, and S. P. S. Beniwal, "Mango malformation," In: Plant Diseases of International Importance (Eds. J. Kumar, H.S. Chaube, U.S. Singh, A.N. Mukhopadhyay), vol. 13, pp. 357-398, 1992. New York, Prentice Hall,pp. 456.

[7] J. Kumar, S. U. Singh, and S. P. S. Beniwal, "Mango malformation: One hundred years of research," *Annual* Review of Phytopathology, vol. 31, pp. 217-232, 1993.

[8] R. Kishun, "Loss in mango fruits due to bacterial canker Xanthomonas campestris pv. Mangiferaeindicae," Proc. 5th Int. Con! Plant Path. Bacteria Cali, pp. 181-184, 1981.

[9] Om. Prakash, K. A. Misra, and A. M. Raoof, "Studies on mango bacterial canker disease," Bio. Memoirs, vol. 20, pp. 95-107, 1994.

[10] E. H. J. Smith, "Bark cracking in mango trees," Citrus and Subtropical Fruit Growers, No. 479. Citrus and Subtropical Fruit Research Institute, Nelspruit (S. Africa), 1973.

[11] P. P. Warne, and R. S. Ganorkar, "Detection of Diseases on Cotton Leaves Using K-Mean Clustering Method," International Research Journal of Engineering and Technology (IRJET), vol. 2(4), pp. 425-431, 2015.

[12] D. Shergill, A. Rana, and H. Singh, "Extraction of rice disease using image processing," International Journal of Engineering Sciences &Research technology, pp. 135-143, 2015.

[13] M. Ranjan, R. M. Weginwar, N. Joshi and A. B. Ingole, "Detection and classification of leaf disease using artificial neural network," International Journal of Technical Research and Applications, vol. 3(3), pp. 331-333, 2015.

[14] R. R. Kajale, "Detection & Recognization of Plant Leaf Diseases Using Image Processing and Android O.S.," International Journal of Engineering Research and General Science, vol. 3(2), pp. 6-9, 2015.

[15] D. Al-Bashish, M. Braik, and S. Bani-Ahmad, "Detection and classification of leaf diseases using K- means-based segmentation and neural-networks-based classification," Inform Technol J, vol. 10(2), pp. 267-275, 2011.

[16] S. N. Ghaiwat, and P. Arora, "Detection and classification of plant leaf diseases using image processing techniques: a review," Int J Recent AdvEngTechnol, vol. 2 (3), pp. 2347-2812, 2014.

[17] A. Garg, V. Vijayaraghavan, S. S. Mahapatra, K. Tai, and H. C. Wong, "Performance evaluation of microbial fuel cell by artificial intelligence methods," Expert Systems with Applications, vol. 41(4), pp. 1389–1399, 2014.

[18] B. S. Ullagaddi, and V. S. Raju, "Disease recognition in Mango crop using modified rotational kernel transform features," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2017.

[19] M. Bhange, and A. H. Hingoliwala, "Smart Farming: Pomegranate Disease Detection Using Image Processing," Procedia Computer Science 58, Elsevier, pp. 280-288, 2017.

[20] S. Kaur, S. Pandey, and S. Goel "Plants Disease Identification and Classification through Leaf Images: A Survey," Archives of Computational Methods in Engineering, Springer, pp. 1-24, 2018.

[21] C. S. Madiwalar, and V. M. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), Pune, pp. 13-18, 2017.

[22] T. Rump, K. A. Mahlein, S. U. Teiner, C. E. Oerke, W. H. Dehne and L. Plumer, "Early detection and Classification of plant diseases with Support Vector Machines based on hyperspectral reflectance," Computers and Electronics in Agriculture, vol. 74(1), pp. 91–99, 2010.

[23] A. Dai, C. Zhang, and H. Wu, "Research of Decision Tree Classification Algorithm in Data Mining," *I*nternational Journal of Database Theory and Application, vol. 9(5), pp. 1-8, 2016.

[24] B. TIhobogang's, and M. Wannous's, "Design of plant disease detection system: A transfer learning approach work in progress," IEEE International Conference on Applied System Innovation, 2018.

[25] H. Kulkarni, and K.R. PatilAshwin, "Leaf Disease Classification Using Artificial Neural Networks and Decision Tree Classifier," Journal of Image Processing & Pattern Recognition Progress, vol. 5(1), 2018.