International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Identification of Plant Leaf Disease using Machine Learning Techniques

Shabari Shedthi B, M Siddappa, Surendra Shetty

Abstract*: — *Plant disease identification and classification is major area of research as majority of people in India depend on agriculture for their main source of income and for food. Identification of the diseases in any crops is challenging since manual identification techniques being used in this are based on the experts advises which may not be efficient. Based on leaf features decisions about variety of diseases are taken. In this paper an automated framework is introduced which can be used to detect and classify the diseases in the leaf accurately. Leaf images are acquired by using digital camera. Pre-processing techniques, segmentation and feature extraction are performed on the acquired images. The features are passed on to the classifiers to classify the diseases. This work has been proposed to classify and distinguish the leaf sample based on its features. The proposed work is carried out with Artificial Neural Network (ANN), Support Vector Machine (SVM) and Naive Bayes classifiers to analyze the result. For given dataset ANN performed better than the other two classifiers.

Keywords*: *Disease identification; Artificial Neural Network; Support Vector Machine; Naive Bayes.

I. INTRODUCTION

Agricultural disease identification is one of the predominant areas of research. Identification of leaf diseases at an early stage will help the farmers to reduce pesticide usage. In recent days lot of research work happened on image processing, which are used to solve the problems in different agriculture applications like plant disease identification [1-2]. Disease of plant can be identified through leaf by using various image processing and machine learning techniques[3-4].

For any image based framework image pre-processing is a basic step, which is a technique to enhance the image quality. The images can be captured from various means like cameras or sensors, space probes etc. In order to ensure the success of the subsequent steps the accuracy of this technique needs to be significantly high, hence dissemination and development of new technology is an important factor in determining the future of agriculture.

An accurate identification of leaf diseases is very important while classifying the leaf diseases and leaf type. Manual classification methods are being used largely by the farmers, which will be done with the expert's advice on identifying the leaf diseases on basis of its geometric parameters. This work proposes a method that processes the captured digital image of leaves and extract the relevant features. To identify the types of diseases texture, shape and color features are used. Image processing techniques are applied to extract various features from leaves and to classify them based on its features. The collected features are then used for disease classification using ANN, SVM and Naive base classifiers.

II. LITERATURE SURVEY

Mishra et.al [5] mentioned and explained many image processing techniques, which are used to detect the leaf diseases. Patel and Joshi [6] discussed about many methods, which is used for classifying the variety of plant diseases.

Sherly et.al[7] reviewed different machine learning algorithms used by many researchers on different types of leaves. Libo Liu & et al [8] used back propagation neural network classifier for differentiating healthy and diseased part of rice leaves in their work.

Sachin et al [9] highlighted different steps and methods involved for plant disease identification i.e. image pre-processing, segmentation, feature vector construction and applying classifier to detect the disease. Paper highlights some segmentation algorithms and feature extraction methods, which can be used for plant disease detection.

Monika et al [10] used image processing to monitor the fruit disease. In this NN with Back Propagation(BP) is used. Three features i.e. color, texture and morphology are extracted and stored in feature vector. They achieved 90% accuracy by considering only morphology over other two. This is done for two apple and three grape diseases.

Suresha et al. [11] presented a method for the identification of diseases in paddy leaves where the leaf images were captured using camera from paddy fields. The RGB images were converted into HSV color space and region of interest is obtained by using Otsu segmentation. Geometric features like area, perimeter, minor axis length and major axis lengths were extracted and classified using K-Nearest Neighbor to achieve 68.1% accuracy. Thilagavathi et al. [12] developed disease diagnosis system for guava leaf. Pre-processing like resizing and image enhancement was done.

Manuscript published on 30 September 2019
* Correspondence Author
Shabari Shedthi B*, CSE,, NMAM Institute of Technology, Nitte,(affiliated to VTU, Belagavi), India. Email: shabarishetty@nitte.edu.in
M siddappa, CSE, SSIT, Tumkur, India.
Surendra Shetty, MCA,, NMAM Institute of Technology, Nitte, (affiliated to VTU, Belagavi), India.

Then using region growing method diseased portion is highlighted. In this CIELAB color space is used and SIFT is used for feature extraction.

SVM and K-NN classifiers were used to classify the diseases, SVM classifier performed better than K-Nearest Neighbor classifier.

Mohammad et al. [13] used SVM classifier to identify cotton leaf diseases i.e with leaf spot and leaf miner disease, K-means clustering is used for clustering.

Prashar et al. [14] used SVM classifier to identify the diseases. Accuracy achieved was greater than 85%.

Singh et al. [15] proposed a methodology to detect the plant leaf diseases. is used. By using Genetic algorithm for segmentation, GLCM for feature extraction and SVM classifier for classification plant disease identifications are done.

Kaushal et al. [16] used K-means for segmentation and K-NN is for plant disease identification. The segmentation is done by using k-means clustering and features are extracted from the images by using grey level co-occurrence matrix..

Rani et al. [17] presented a method to identify the pests in the leaf images. Contrast enhancement was done enhance image and using k-means algorithm segmentation is applied. Features are extracted from theses image then SVM classifier is applied for identification. Percentage of affected regions was calculated.

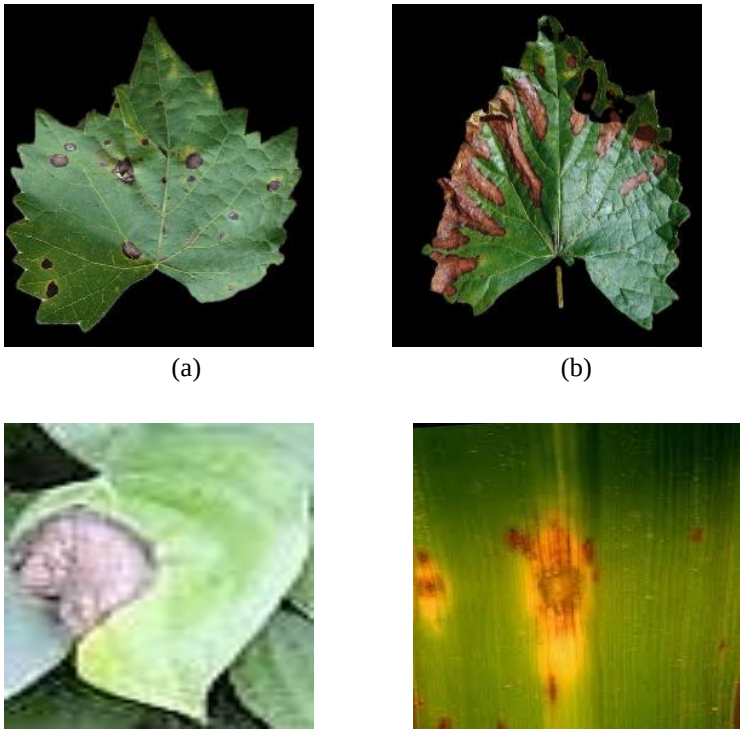
Mondal et al. [18] presented a method to detect yellow mosaic virus disease where the leaf images were captured using camera of high resolution. The input images in RGB color space were converted into grayscale image and morphological operations are applied on them. The features including mean, standard deviation, entropy, homogeneity, mode and median are extracted. Naive Bayes classifier was used to classify the diseases.

III. PROPOSED METHODOLOGY

An architecture of proposed leaf disease classification system is shown in Fig. 2. The aim of the paper is to develop a leaf disease identification and classification system using its color, texture and shape features, which classifies the type of disease as well as the portion of the leaf affected by the disease using ANN, SVM and NB classifiers. It is carried out with the following steps,

A. Image Acquisition

Input image samples are containing 120 leaf images are collected and stored. Collection of the leaf images are done using digital camera and internet sources. Collected leaf sample contains the infected Grape leaves with Black Rot and Black Measels diseases and Quick Wilt diseased leaves of pepper, Rice Blast disease of paddy and also the healthy leaves of above spices. Input images are shown in Figure. 1.



(a) (b) (c) (d)

Fig .1. Diseased plant leaf images.

B. Pre-processing

To improve the image quality, suppress the undesired destortion and for noise removal, pre-processing is done on the gathered images. The input images will be of different sizes which are initially resized to 256x256 pixels to maintain uniformity. Median filtering is applied to smoothen the images by removing the noise in the image.

C. Segmentation

Segmentation is done by using K-means clustering to obtain the interested region of the leaf. K-means clustering technique with 'k' value as *three* is used to segment the images based on the color to obtain the portion of diseased leaf area. The diseased area will be present in any of the three clusters, which need to be selected for the next phase.

D. Feature Extraction

It is a type of dimensionality reduction technique and as a compact feature vector the interesting parts of an image are represented efficiently. From the segmented image shape, texture and color features are extracted.

Color features like color moments: Mean, Standard Deviation is extracted. Using Grey Level Co-occurrence matrix (GLCM) texture features including Homogeneity, Mean, Standard deviation, Variance and Entropy are extracted. By using the function regionprops(), shape features of leaf are extracted. The shape features extracted are: Area, Perimeter and Eccentricity.

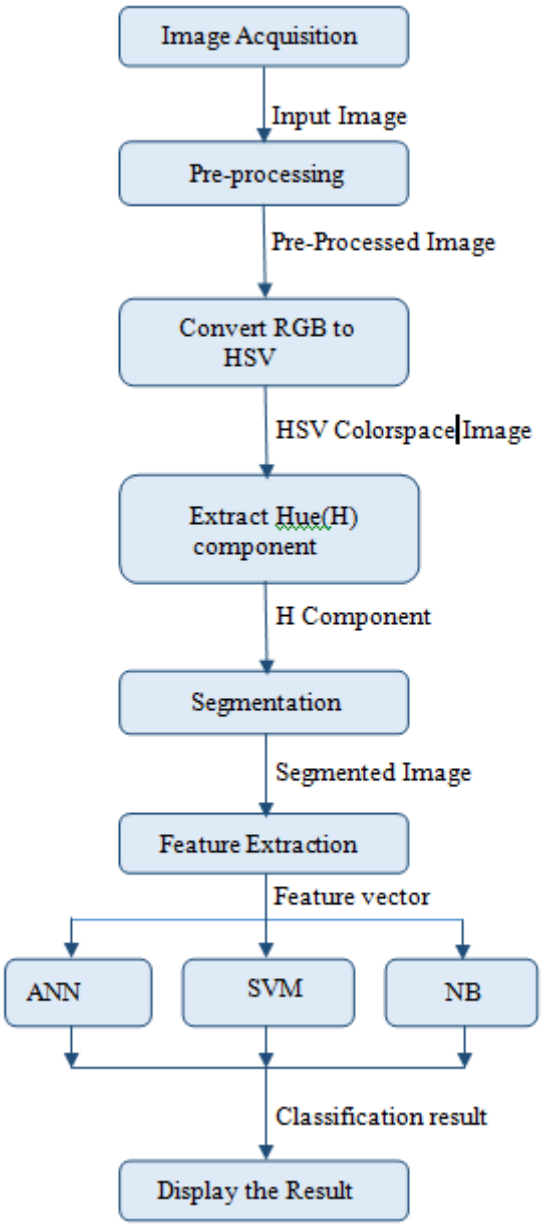


Fig .2. Classifiaction system

E. Classification

Classification is used to identify the class of the new observation using training data whose class label is known. It is performed by using three different machine learning techniques. For classification Artificial Neural Network, Support vector machine and Naive Bayes classifiers are used. These are supervised learning based classifiers, which makes predictions using known datasets are described below:

Artificial Neural Network (ANN): This concept is derived fropm biological human neurons system. ANN is a mathematical model used for pattern recognization. In this each node behave like neuron. There are different type of ANN is there based on its network structure. In that Back propgagation neural(BPN) networkis one and which is widely used. In BPN error is back propagated from output layer to input layer for correction. It can handle noisy data effectively and well suited for complex problems.

Support Vector Machine (SVM): Support vector machines can have good prediction speed and memory usage with few support vectors. Best hyper-plane is found which separates one class of data points from the data points of other class. The hyper-plane which has the maximum margin between the support vectors of two classes is known as the best hyper plane. The slab containing maximal width which is parallel to the hyper plane is referred to as the margin.

Naive Bayes(NB) Classifier: Based on Bayes theorem classification is performed using probabilistic analysis, which is a statistical classifier.

F. Performance evaluation

The classifiers will produce the, disease type as well as the portion of disease affected area in terms of percentage. The performance evaluation of the obtained result is done based on the values of accuracy.

Accuracy: It is the ratio of obervation predicted correctly to the total number of observations. It is calculated as,

$$\text{Accuracy} = \frac{\text{Correctly predicted observations}}{\text{Total number of observations}} \quad (1)$$

Percentage of disease effected to the leaf is calculated using Equation (2). This will give the severity of the disease. This information will help the farmers to decide or select amount of pesticide to be used.

$$\text{Pecentage of disease} = \frac{\text{Pixels in disease portion}}{\text{Total pixels in leaf area}} \quad (2)$$

IV. RESULT AND DISCUSSION

Acquired 120 images are stored in database and Median filter is used for noise reduction. Images are segmented using K-means algorithm to form clusters containing diseased portion of the leaf which is shown in Fig. 3 and cluster with diseased portion is selected and their features are extracted.

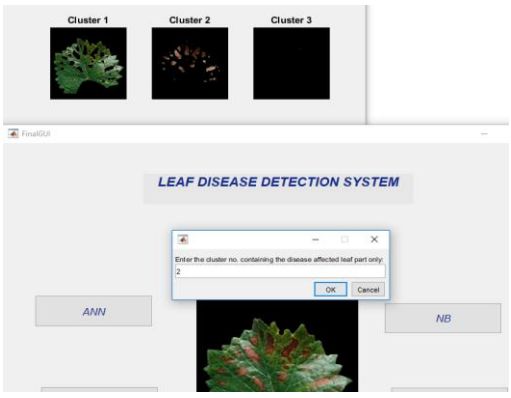


Fig .3. Segmented images of input leaf image.

Extracted features are stored in feature vector and given as input to classifier. Classifier is trained with respective target values in training phase. The model is trained using ANN, SVM and NB classifiers.

Test image is pre-processed, segmented, features are extracted and trained classifiers are used to obtain the type of the leaf disease. Output of the classifiers is shown in Fig. 4.

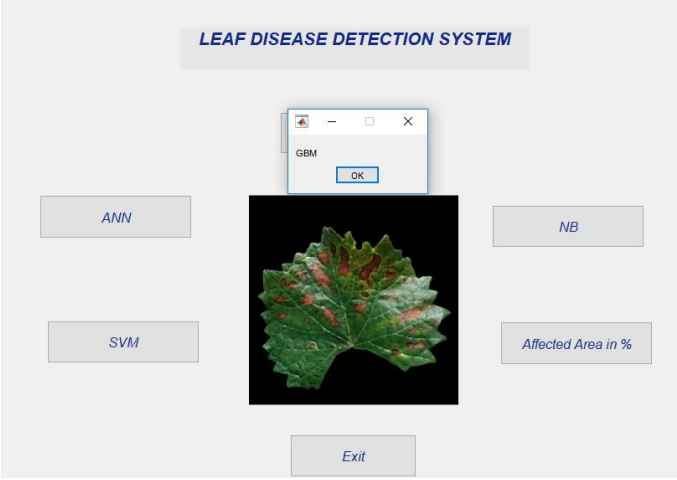


Fig. 4. Leaf disease classification result

The portion of leaf affected by the disease is calculated in terms of percentage as shown in Fig. 5.

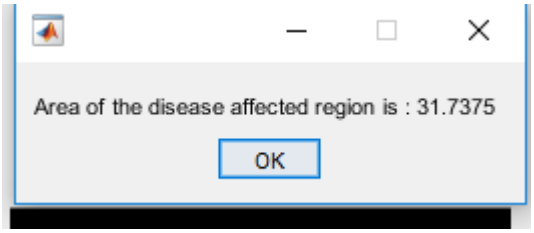


Fig. 5. Percentage of disease

Performance evaluations of KNN, SVM and NB classifiers are done based on Accuracy which is show in Figure 6.

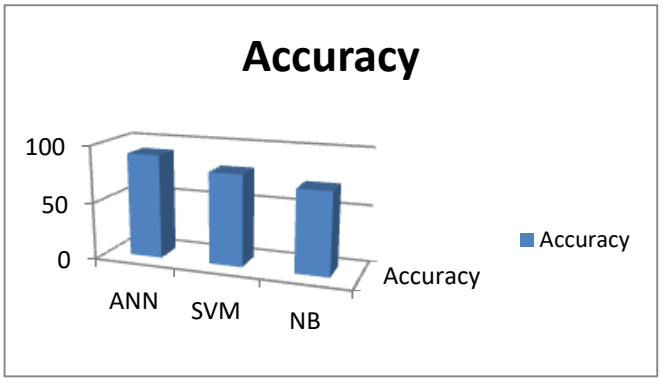


Fig. 6. Performance evaluation of classifiers

V. CONCLUSION

The developed model is used for identifying and classifying the leaf diseases which also gives affected portion of the leaf in percentage. This helps to know the severity of disease and through this, framers can take control measures to save the plant. The proposed work presents the classification accuracies for diseased leaves obtained through image processing by using ANN, SVM and NB classifiers. In this work ANN has achieved higher accuracy. In future, hybrid algorithm can be used for improving the recognition rate of plant disease with more dataset.

REFERENCES

1. S.B. Dhaygude, N.P.Kumbhar, "Agricultural plant leaf disease detection using image processing," International. J. Advanced. Res. Electr. Electron. Instrum. Eng., vol. 2, pp. 599-602, Jan. 2013.
2. G. Savita, A. Parul, "Detection and classification of plant leaf diseases using image processing techniques: a review," International J. Recent Adv. Eng. Technol., vol. 2, pp. 2347–812, 2014.
3. S.R. Vaidya, "Leaf Disease Detection using Image Processing", Volume 5, International Conference on recent trends in Engineering Science and Technology, 2017, pp. 333-336.
4. S. B. Patil, and S. K. Bodhe, "Leaf disease severity measurement using image processing," Int. J. Eng. Technol., vol. 3, pp. 297–301, Oct. 2011
5. Mishra, Bharat, et al. "Recent technologies of leaf disease detection using image processing approach—A review." *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. IEEE, 2017.
6. Patel, Arpita, and Mrs Barkha Joshi. "A Survey on the Plant Leaf Disease Detection Techniques.", International Journal of Advanced Research in Computer and Communication Engineering(IJARCCE), Vol. 6, Issue 1, January 2017.
7. Annabel, L. Sherly Puspha, T. Annapoorani, and P. Deepalakshmi. "Machine Learning for Plant Leaf Disease Detection and Classification–A Review." *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2019.
8. Liu, Libo and Guomin Zhou Guomin Zhou, *"Extraction of the Rice Leaf Disease Image Based on BP Neural Network." International Conference on Computational Intelligence and Software Engineering, 1-3, 2009*
9. P. Sachin d. Khirade, "plant disease detection using image processing," IEEE international conference on computing communication control and automation, pune, 2015.
10. K. B. Monika jhuria, "image processing for smart farming: detection of disease and fruit grading," in proceedings of the IEEE second international conference on image information processing , 2013.
11. Suresha, M., Shreekanth, K. N., & Thirumalesh, B. V.. Recognition of diseases in paddy leaves using knn classifier. In *Convergence in Technology (I2CT), second International Conference for* (pp. 663-666). IEEE, 2017.
12. M. Thilagavathi, S. Abirami, "Application of image processing in diagnosing guava leaf diseases", International Journal of Scientific Research and Management (IJSRM), Volume 5, Issue 07, July 2017.
13. Mohammad, M. B., Srujana, R. N., Jyothi, A. J. N., and Sundari, P. B. T., "Disease identification in plants using K-means clustering and gray scale matrices with SVM classifier", International Journal of Applied Sciences, Engineering and Management, Volume 05, No. 02, March 2016.
14. Prashar, K., Dr. Rajneesh Talwar, Dr. Chander Kant, "Robust Automatic Cotton Crop Disease Recognition (ACDR) Method using the Hybrid Feature Descriptor with SVM". 4th International Conference on Computing on sustainable Global Development, March, 2017.
15. Singh, V., Misra, A. K.: Detection of plant leaf diseases using image segmentation and soft computing techniques. Information Processing in Agriculture, 4(1), 41-49., 2017.
16. Kaushal, G., Rajani, Bala, "GLCM and KNN based Algorithm for Plant Disease Detection", International Journal of Research in Electrical, Electronics and Instrumentation Engineering, Volume 6, Issue 7, July 2017.
17. Rani, R. U., and Amsini, P., "Pest identification in leaf Images using SVM classifier", International Journal of Computational Intelligence and Informatics, Volume 6, No. 1, June 2016.

18. Mondal, D., Chakraborty, A., Kole, D. K., & Majumder, D. D. (2015, October). Detection and classification technique of yellow vein mosaic virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier. In Soft Computing Techniques and Implementations (ICSCTI), 2015 International Conference on (pp. 166-171). IEEE.

AUTHORS PROFILE

Shabari Shedthi B, working as Assistant Professor in the Department of CSE, NMAM Institute of Technology, Nitte. She has completed her M.Tech in Computer Science and Engineering from Visvesvaraya Technological University in the year 2013.. She has 6 years of teaching experience and 3 years of research experience. She has published many research papers in different international journals and conferences. . The Research areas of interest are of Image Processing and Pattern Recognition

Dr. M Siddappa is working as Professor and Head in department of computer science & Engg., SSIT Tumkur, Karnataka, India. He has received PhD in computer science and Engineering from MJR University. He has published many research papers in international journals and presented several papers in international conferences. The Research areas of interest are Data Structure, Image Processing, Networking, Pattern Recognition.

Dr. Surendra Shetty, Professor & Head, Dept. Of MCA, had completed his B.Sc. in 2001 and Master of Computer Applications during 2004. Dr. Surendra Shetty had been awarded his doctoral degree for his research work “Audio Data Mining Using Machine Learning Techniques” in 2013 from university of Mangalore. He has published more than 25 research papers in different international journals and conferences. The Research areas of interest are Cryptography, Data mining, Pattern Recognition, Speech Recognition, MIS, Software Engineering and Testing.

.