International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3S, February 2019

Plant Disease Identification and Classification using Image Processing

E.Vamsidhar, P. Jhansi Rani, K. Rajesh Babu

Abstract: ***An agricultural sector plays a vital role in the economy of country. Agricultural output is very vital in many developing countries. Increase in population and increase in the life expectancy is pressurizing the agricultural sector to come out with new types of high yielding crops. The diseases in the plants are common, early detection and controlling increases the yield of a crop. Development of technology in the field of computer science can be applied to detect these diseases early. Image processing and classification methods can be applied to identify the plant disease in the early stage. This paper developed a segmentation technique for automatic detection and classification of plant leaf diseases. Features are extracted and selected features are used for training and support vector machine (SVM) and artificial neural network (ANN) classifiers. The results obtained are satisfactory.***

Key words: Image processing, CBIR, SVM, Plant Disease, ANN.

I. INTRODUCTION:

Many of the developing countries are based on the agricultural sector. More than 60 % of the population depend on this sector. Agricultural section is the segment of an economy related with the improvement of plants and raising of animals and fish for monetary and residential uses. The farming area is apparently the most vital fragment of the economy as it is imperative for human presence through the generation of sustenance, accommodation through wood, and sartorial. Agribusiness was one of the features in human advancement and was a factor prompting the ascent of stationary human development [1]. Most of the African countries and some Asian countries like Burma, Thailand etc have major dependency on the agricultural sector [2]. The climate has a direct effect on the agriculture and fisheries. The build of carbon dioxide in the atmosphere can intensify some harvest yields in a couple of spots, despite the fact that it requires to comprehend these points of interest, soil dampness, water accessibility, and distinctive conditions ought to moreover be met. Changes in the recurrence and seriousness of dry spells and surges could present difficulties for agriculturists and farmers and debilitate sustenance security. In general, climate change makes it difficult to grow crops in the way it is done at similar spots from that they done before. Climate change is also a factor along with other advancing components that influence agricultural production, like changes in cultivating practices and modernization [3]. The effect of the heat growth in the climate has led to the build-up of the pests and disease that is affecting the flowering in the plants. As quoted by Shri, B. Venkateshwarlu, former director at International Central Research Institute for Dryland Agriculture (CRIDA), Hyderabad, said that "Changes in climate influences availability, access and absorption. The decrease in production makes the less availability of food grains which have a large influence mostly on poor people where they cannot afford to purchase the food which in turn effects their health".[4] The agricultural production at the global stage has been effected by the plant diseases as the combined effect of the pests and diseases on plants can end up to loss of over 50% in the major crop,20% in case of prime food and cash crops[5]. This has a large effect on the income of the farmer and thereby it also effect the GDP of the countries. The growth in technology has raised the agricultural output and also able to feed the rising populations. In the past the leaf samples are taken to the local agricultural centres to recognize the disease. The farmers and the agrarian professionals visually screen the crops to find out that any attack has taken place on these crops and plants and classify them. This practice is more monotonous, overwhelming and subjective. The decision making capability of the human being depends on many factors. The ability to judge be influenced by on the physical conditions such as exhaustion, vision conceptual state stress surroundings, climate improper lighting etc. The growth of internet and the rise of the use of the smart phones having high computing capabilities and high resolution cameras now can be utilized to identify the diseases in a shorter time [6].The current technique in support of plant infection detection is essentially bare eye perception by specialists in the course of which recognition and location of crop infections is done. A huge group of specialists and in addition nonstop checking of plant is necessary, creating a huge costs especially dealing with huge ranches. In the meantime, in a few nations, ranchers don't have legitimate offices or even thought that they can con-affability to specialists. Because of which counselling specialists even cost high and additionally tedious as well. In such conditions, the proposed system turns out to be helpful in checking huge fields of yields. Programmed recognition of the sicknesses by simply observing the side effects on the plant leaves makes it less demanding and less expensive.

Manuscript published on 28 February 2019.
***** Correspondence Author (s)
Dr.E.Vamsidhar, Assoc.professor, Department of CSE, KL University, Vaddeswaram, A.P, India. (enireddy.vamsidhar@gmail.com)
P.Jhansi Rani, Assoc.professor, Department of CSE, KL University, Vaddeswaram, A.P, India (jhansi.infoarea@gmail.com)
K.Rajesh Babu, Assoc.professor, Department of ECE, KL University, Vaddeswaram, A.P, India (krajeshbabu@kluniversity.in)

This likewise bolsters machine vision to professional vide picture based programmed process control, review, and robot control [7, 8, 9].

It is a recognizable proof by ocular way, more relentless assignment and in the meantime, less precise and should be possible just in constrained territories. While if programmed identification system is utilized it will take less endeavours, less time and turn out to be more exact. In crops, a number of broad maladies is visualized as dark coloured and pale spots, premature and tardy sear, and some are parasitic, bacterial and viral infections. Image Processing is utilized for estimating influenced territory of malady and to decide the distinction in the shade of influenced region [10, 11]. Segmentation is a procedure to group an image into diverse portions. Humans can simply identify, distinct and interpret the specific objects whereas the processing machines do not have such intelligence in identifying these objects, hence many different methods have been developed for the segmentation [12]. The plant diseases can mainly categorized as fungal and Bacteria [23].

II. MATERIALS AND METHODS

The images of the leaves are obtained from different plants and having different defects. Using the image pre-processing techniques the affected areas are identified in them. The features are obtained from the leaves to analyze them in the later steps.

The step by step process of the approach for this process had been given below

1. The image is acquired using a digital devices.
2. The pre-processing of the image deals with many issues like removing of the noise or distortion in image, enchasing quality of image, cropping the image to find the interested area on it and finally utilizing smoothing filter to smooth down the image.
3. As the images are green since they are leaves we can mask the green pixels are they are healthy areas. Threshold value is calculated that can be used for those pels. Each pixel is compared with the threshold value thereby the green pixels are masked and the other colour components blue, red, green are consigned to zero value.
4. The infected regions are identified and these regions the boundaries are identified and masked cells are removed.
5. Acquiring constructive segments to categorize the leaf diseases.
6. Apply genetic algorithm to segment the image.
7. Calculate the features by means of color co-occurrence method

Hierarchical clustering a basic clustering method to build a hierarchy of clusters. The merge or split execution has an impact on the quality on this method [13].K-means clustering is an unsupervised widely used and widely used clustering technique [14]. In this clustering both hierarchal and k-means are combined to obtain a better results [15]. The performance of the K-means clustering depends on the good choice of the k having less number of cycles for it converge. Any wrong choice decreases the efficiency of it. In [16] author presented many limitations of the k-means clustering algorithm. To improve further the clustering is further optimized using the genetic algorithm as it is easy to understand, provision for multi-objective and better for “noisy” environments. Initial population is formed from randomly generated several individual solutions. Optimal solution is formed from them by scattering them in various regions. The individual solutions are identified by using the fitness function [15]. Individual chromosome is a solution for the sequence of cluster centres. The method builds k-Mean constructed significantly high quality clusters. The examination has introduced a diminished list of capabilities based methodology for acknowledgment and order of pictures of plant diseases. The outcomes uncover that SVM classifier is more appropriate for distinguishing proof and arrangement of plant infections influencing agribusiness/cultivation crops Feature extraction is one of the important steps in the image processing. The color co-occurrence method is one of the method to find the features. The texture and color features are mostly used as features and they are utilized in this process. RGB images are converted into the HIS color space representation. The author in [17] have completely defined 14 measurement feature measures that can be computed from the co-occurrence network with the goal to depict the texture of the images. The following feature measurements defined below are calculated as[18]

(i) Contrast is a proportion of powers differentiate between a pel and its neighbor. For a "consistent" picture contrast is NIL.

$$cont = \sum_a \sum_b |a-b|^2 p(a,b)$$

(ii) Local homogeneity estimates proximity of circulation of components in GLCM to its diagonal, generally it is 1.

$$homog = \sum_a \sum_b \frac{1}{1+|a-b|^2} p(a,b)$$

(iii) Correlation is a proportion of how corresponded a pixel is to its neighbor over the entire picture.

$$corr = \sum_a \sum_b \frac{(a-\mu_i)(b-\mu_j)p(a,b)}{\sigma_a \sigma_b}$$

(iv) Cluster shade and prominence characterizes the affinity of clustering of pixels in the ROI.

$$cluster\ shade = \sum_a \sum_b (a+b-\mu_a - \mu_b)^3 p(a,b)$$

$$cluster\ prominence = \sum_a \sum_b (a+b-\mu_a - \mu_b)^4 p(a,b)$$

(v) Entropy is a statistical measure to characterizes the texture of an image

$$Entropy = \sum_a \sum_b p(a,b) \log p(a,b)$$

(vi) Dissimilarity evaluate distance between pairs of objects in region of interest.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$Dissimilarity = \sum_{a}\sum_{b}|a-b|p(a,b)$$

(vii) Maximum probability measures the maximum likelihood of producing the pixels of interest.

$$Max.prob = maxp(a,b)\forall(a,b)$$

(viii) Energy provides sum of squared elements in the GLCM. It has values between 0 and 1

$$Energy = \sum_{a,b} p(a,b)^2$$

Choice of an arrangement of fitting information include factors is a vital issue in the working of a classifier.

The reason for feature subset collection is to locate littlest arrangement of highlights with the intention of getting the better prescient execution. As a result of the scourge of dimensionality, usually fundamental and valuable to constrain the quantity of information includes in a classifier is to get better accuracy with less computation. Various element choice techniques have been produced in the example acknowledgment writing [19]. In this work the best discriminating feature is selected by analyzing the classes' separability power of each feature.

III. SUPPORT VECTOR MACHINE (SVM)

SVM is a widely used supervised learning classification algorithm [20]. In general feature vectors and class labels are available in the datasets then SVM (supervised) otherwise labels can use unsupervised clusters.SVM can be developed using different kernel functions, as non-homogenous function homogeneous polynomial, and radial bias function (RBF).

To classify the images using SVM is obtained by taking into account the training feature vector comparing with the testing feature vector to conclude target class, as

$$P = \{(x,y)|x_i \in R^n, y_i \in \{-1,1\}\}t_1 = 1)$$

The SVM is used for the classification of the diseaes which are classified by using the selected features of the texture and color.

IV. ARTIFICIAL NEURAL NETWORKS (ANNS)

Artificial Neural Networks can be utilized for characterization, enhancement and forecast issues and have generally connected in various fields. These can be also used in the farming applications[21]. The feed forward back propagation neural network has three layers input, hidden, output layers which utilizes a back propagation algorithm is most popular in ANN and the features can be given as input to the weights[22]. Utilizing numerical capacities, information are prepared inside each layer containing hubs and the handled data is given to the resulting layers. This is an iterative model where association weights are balanced so that the distinction in assessed and genuine yields is lessened, by over and over displaying forecasts and reactions to the systems. To improve the performance of the network the weights are optimized using the genetic algorithm.

V. RESULTS AND DISCUSSIONS

The database of the images are created where as Alternaria alternate (fungal), Anthracnose, Bacterial Blight (bacteria), Cercospora Leaf Spot, Bacterial leaf spot, frog eye leaf spot, sun burn disease and healthy leaves images are also considered. The images are segmented using hybrid k-means algorithm, where hierarchal clustering is applied then the results are given as input to the k-means where the genetic algorithm is used to optimize the cluster for better performance. In this obtained data is divided such that 80% for training and 20% for testing.

The texture and colour features are extracted and these features are selected to get the better feature set as input to the classification algorithms. The k-means can also be used as classification and hence it is used as classifier and found out that the accuracy presented by it 85.3%.

Initially input to the SVM is given without feature selection and using k-means for segmentation of the images. The SVM used for the classification and it obtained 90% accuracy. The Linear Kernel is applied in SVM and found out that classification accuracy is 89%.The RBF kernel is considered 88.8% and SVM polynomial kernel gave 90.2%.

As SVM effectiveness depends on the selection of kernel, Kernel parameters and soft margin parameter which can maximize the efficiency, reduce error and over fitting. Now the proposed segmentation process is applied and features are obatined from it and best feature sub set is used as input to the SVM and classification accuracy obtained 95.63%, 94.23% and 95.78%.

The other classification algorithms are used for comparision C4.5 89% and Naive Bayes 86%. The MLP with the Back-propagation as the learning algorithm is used and accuracy obtained 88.59%. Now the optimized MLP have obtained accuracy of 91.45%.

Table.1 Showing overall accuracy, precision, recall and F-measure

S.No	Classifier	Precision	Recall	F-Measure	Classification Accuracy
1	K-means	0.823	0.846	0.843	85.3
2	C 4.5	0.882	0.874	0.891	89
3	Naive Bayes	0.850	0.854	0.842	86
4	MLP	0.823	0.847	0.834	88.59
5	MLP with Genetic Optimization	0.88	0.896	0.912	91.45
6	SVM(Linear Kernel)	0.921	0.954	0.946	95.63
7	SVM(RBF Kernel)	0.914	0.948	0.935	94.23
8	SVM(Polynomial Kernel)	0.928	0.959	0.951	95.87

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

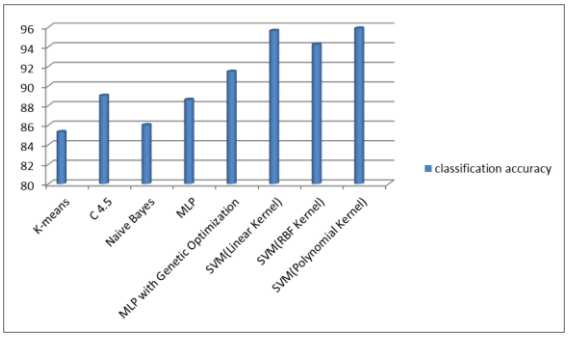

Fig.12 Showing the classification accuracy

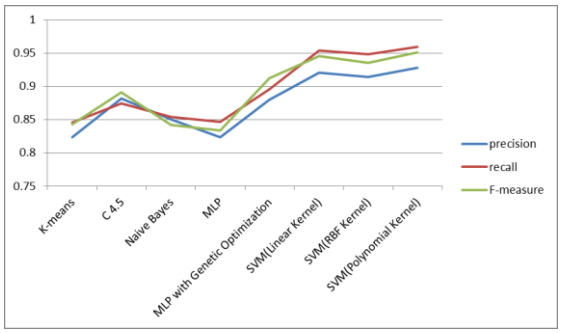

Fig.13. Showing the Precision, recall and F-measure

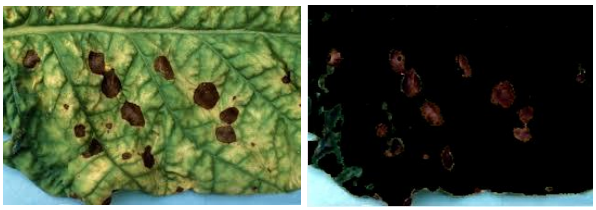

Fig.2. Information along with segmented image of leaf and disease is Alternaria Alternata.

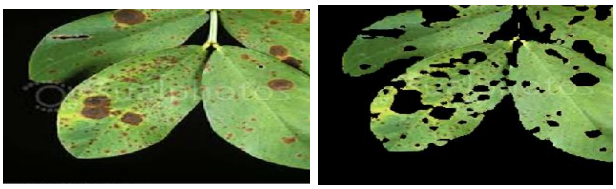

Fig.3. Information along with segmented image of leaf and disease is Anthracnose

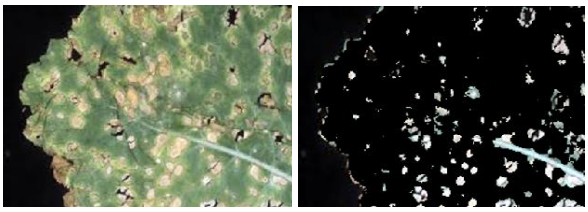

Fig.4. Information along with segmented image of leaf and disease is Bacterial Blight

Fig.5. Information along with segmented image of leaf and disease is Cercospora Leaf Spot

Fig.6. Information along with segmented image of a healthy leaf

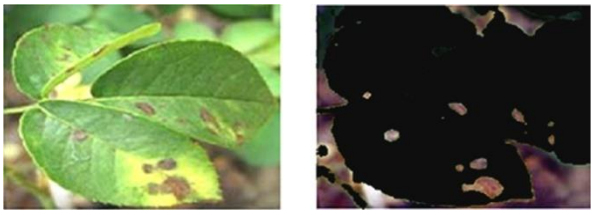

Fig. 7 – Information along with segmented image of leaf and disease is bacterial leaf spot.

Fig. 8 – Information along with segmented image of leaf and disease is early scorch disease.

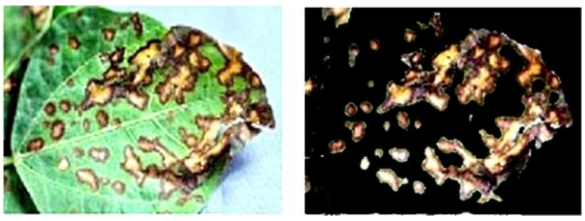

Fig. 9 – Information along with segmented image of leaf and disease is bacterial leaf spot.

Fig. 10 – Information along with segmented image of leaf and disease is sun burn disease.

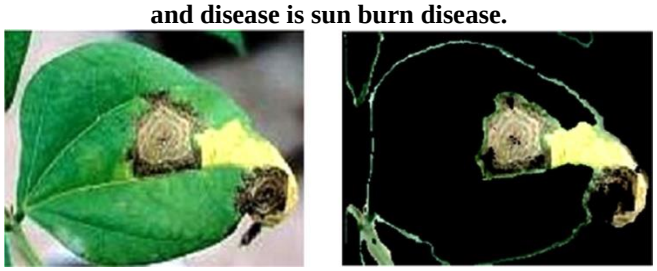

Fig. 11. Information along with segmented image of leaf and disease is fungal disease.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

VI. CONCLUSION:

In this work a novel K-means clustering is designed to obtain the segmentation of the leaf images and then the features are obtained from it. The features are selected and selected sub-set is used for classification of various diseases using different classification methods. The training set is used for training and training set is used for testing the algorithms. It has been found that the SVM had shown better results in identification and classification of fungal diseases on cereal crops whereas NN has shown better classification in fungal disease identification in case of the vegetable crops. The results shown here are overall accuracies obtained from the different diseases and different crops and diseases. Early detection of the disease can prevent the huge loss for the farmer and thus the productivity increases and there by the economy also increases. The advancement of drone technology where the data is collected and this data can be analyzed with these type of systems easily and also this information can be utilized for spraying the pesticides depending on the severity of the diseases selectively by using drones itself so that time, effects on the humans can also be reduced. Further deep learning algorithms can be used to train the network over a large data set there there by a single classifier can obtain a better identification and classification of the different disease over a wide variety of plants.

REFERENCES:

1. http://unctad.org/en/pages/newsdetails.aspx?OriginalVersionID=1581
2. https://www.worldatlas.com/articles/countries-most-dependent-on-agriculture.html
3. USGCRP (2014). Ziska, L., A. Crimmins, A. Auclair, S. DeGrasse, J.F. Garofalo, A.S. Khan, I. Loladze, A.A. Pérez de León, A. Showler, J. Thurston, and I. Walls, 2016: *Ch. 7: Food Safety, Nutrition, and Distribution*. The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. U.S. Global Change Research Program, Washington, DC, 189–216.
4. https://www.downtoearth.org.in/news/climate-change-causes-about-1-5-per-cent-loss-in-india-s-gdp-57883
5. H. R. Gautam, M. L. Bhardwaj and Rohitashw Kumar "Climate change and its impact on plant diseases" Vol. 105, No. 12 (25 December 2013), pp. 1685-1691
6. ITU (2015). ICT Facts and Figures – the World in 2015. Geneva: International Telecommunication Union
7. Dhaygude Sanjay B, Kumbhar Nitin P. Agricultural plant leaf disease detection using image processing. Int J Adv Res Electr Electron Instrum Eng 2013;2(1)
8. Arivazhagan S, Newlin Shebiah R, Ananthi S, Vishnu Varthini classification of plant leaf diseases using texture features. Agric Eng Int CIGR 2013;15(1):211–7.
9. Kulkarni Anand H, Ashwin Patil RK. Applying image processing technique to detect plant diseases. Int J Mod Eng Res 2012;2(5):3661–4.
10. Ghaiwat Savita N, Arora Parul. Detection and classification of plant leaf diseases using image processing techniques: a review. Int J Recent Adv Eng Technol 2014;2(3):2347–812. ISSN (Online).
11. Bashir Sabah, Sharma Navdeep. Remote area plant disease detection using image processing. IOSR J Electron Commun Eng 2012;2(6):31–4. ISSN: 2278-2834.
12. Beucher S, Meyer F. The morphological approach to segmentation: the watershed transforms. In: Dougherty ER, editor. Mathematical morphology image processing, vol. 12 New York: Marcel Dekker; 1993. p. 433–81.
13. V.S.V.S. MURTHY, E.VAMSIDHAR, J.N.V.R. SWARUP KUMAR, P.SANKARA RAO Content Based Image Retrieval using Hierarchical and K-Means Clustering Techniques International Journal of Engineering Science and Technology Vol. 2(3), 2010, 209-212.
14. MacQueen, J. (1967) Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.
15. Koza J., Forest B., David Andre and Martin Keane. 1999. Genetic Programming III: Darwinian Invention and Problem Solving. Morgan Kaufmann Publishers.
16. X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.H. Zhou, M. Steinbach, D.J. Hand, D. Steinberg, "Top 10 algorithms in data mining", Knowledge and Information Systems, January 2008, Volume 14, Issue 1, pp 1-37.
17. Robert M. Haralick, "Statistical and structural approaches to texture.", Proc. IEEE, vol. 67, no. 5, pp. 786 - 804, 1979.
18. Aborisade, D.O., Ojo, J. A., Amole, A.O., Durodola A.O. "Comparative Analysis of Textural Features Derived from GLCM for Ultrasound Liver Image Classification" International Journal of Computer Trends and Technology (IJCTT) – volume 11 number 6 – May 2014.
19. K. Funahashi," Introduction to Statistical Pattern Recognition". New York: Academic, 1990.
20. Vapnik V, Vapnik VN, (1998) Statistical learning theory, vol 1. Wiley, New York.
21. Tang L, Tian L, Steward BL (2003) Classification of Broadleaf and Grass Weeds Using Gabor Wavelets and An Artificial Neural Network. Trans ASAE 46(4):1247–1254.
22. Monica Jhuria, Ashwani Kumar, and Rushikesh Borse, "Image Processing For Smart Farming: Detection Of Disease And Fruit Grading", Proceedings of the 2013 IEEE Second International Conference on Image Information Processing ICIIP-2013.
23. Pooja Kulinavar and Vidya I. Hadimani Classification of leaf disease based on Multiclass SVM classifier International Journal of Advance Research, Ideas and Innovations in Technology Vol. 3,issue.4,pp 321-327.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)