Automatic Detection of Tomato Leaf Deficiency using Soft Computing Technique

S. Sivagami , S. Mohanapriya

***Abstract:** Indian Economy mostly depends on Agriculture. Agriculture and its value-added products will occupy considerable amount of gross domestic product (GDP) and provides employment to more than half of the country's workforce. Among all the countries India is one of the world's largest producer of agriculture and horticulture crops. Among all the vegetables Tomato is one of the most important vegetable used to consume all over the world. Disease easily affect the tomato plant due to insects and nutrient deficiency. To detect nutrient deficiency using image segmentation and classification is the main focus of this paper. If detect nutrient deficiency in early stage then he yields increased and the disease caused due to lack of nutrient deficiency also reduced. In this paper k-means and Expectation maximization segmentation algorithms are used for segmentation and SVM classifier used for classification. Based on the results Expectation Maximation provide better result than K-means segmentation.*

***Keywords:** deficiency detection, k-means, Expectation-Maximization, SVM.*

I. INTRODUCTION

Agriculture is mother of all culture it plays a vital role in the development of human civilization. It is required to increase the crop yield to fulfill the food requirement of large population. Most of the income is coming from agriculture only in India and the government has spent lot of money for new techniques and methods to improve yield from agriculture. Not only to increase the yield but also to improve the quality. The economic status of farmers will definitely improve when good quality and high yield fruits are harvested and sold to the market as well as fruit industry. Good quality fruits contain enough nutrient to the consumers. Therefore, in field of agriculture, early detection of deficiency and disease in plants plays a vital role. Automatic deficiency detection in initial stage is very useful for the farmers to supplement the required nutrient to avoid growth related problems. Till now the deficiency is identified by the expert through naked eye observation. In general, this kind of deficiency is identified through experienced people but due to environment changes the manual prediction becomes too difficult. Automatic detection of nutrient deficiency by seeing the symptoms on the leaves and stems makes it easier as well as cheaper. Natural factors like soil type, nutrients available, water and seed quality etc., affects the productivity and quality of the cultivated fruits and vegetables. Sometimes it is necessary to supplement extra nutrients if natural resources do not fulfil the crop requirement. For this extra supplement of nutrients, the productivity cost will increase. By utilizing technology development, we can use image processing for detection of plant disease by analysing its deficiency. Generally, we can observe the symptoms of deficiency on leaves, and stems. so here we use leaves for identification of deficiency in affected plants. Nitrogen, phosphorus and potassium are primary nutrients and very important to tomatoes Plants to grow. Nitrogen helps growth and chlorophyll production, and phosphorus helps tomatoes grow and cope with stress while aiding in energy production. Potassium is also important nutrient it provides immune to the plant to fights against disease, Potassium also improves the quality tomato and it regulates CO_2 intake so helps in photosynthesis. Tomatoes also need the "secondary nutrients," calcium, magnesium and sulfur, but in lesser amounts than the primary nutrients. Calcium improves is the most important secondary nutrient, it also provides immunity to the plant and main benefit of soil is it allow water penetration to the plant. Photosynthesis and chlorophyll both rely on magnesium, which helps their overall quality. Deficiency in sulfur will produce yellow leaves. Tomatoes also need little number of micronutrients too, the micro nutrients are Zinc, iron, boron, chloride, molybdenum, copper and manganese. Molybdenum, helps to efficiently use nitrogen, while zinc helps regulate growth and promotes proper sugar consumption. Boron helps the development of good quality fruit and seeds. Nutrients are observed from soil for growth of plants. the growth of plant/crop affected by deficiency of these nutrients. Therefore, analysing nutrients status will play crucial role in agriculture and farming. The deficiency of nutrients in plants will reflect in its leaf. These symptoms are listed as follows, colour of the leaf's changes in to yellow colour, curled edges of the leaves, size of the leaf reduced to smaller size interveinal chlorosis, necrosis, distorted edges, One critical issue in identifying deficiency is similar symptoms present in old and younger leaf does not have same nutrient deficiency. Image processing is widely used in agriculture sector and is advancing with new techniques. Some of the applications are fruit grading, weed detection, crop detection, crop management, disease identification, deficiency identification etc. Image segmentation means divide or partition the image into various parts called segments.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author
S. Sivagami*, Assistant Professor, Department of Computer Science, Adhiyamaan College of Agriculture and Research, Hosur, Tamilnadu.
Dr. S. Mohanapriya, Head of the Department, Department of Computer Science, KSR College of Arts and Science for Women, Tiruchengode, Tamilnadu.

It's waste of time to process the entire image at the same time as there will be regions in the image which do not contain any information. The image is divided into lots of segments with the help of image processing algorithms with the rule that relationship between pixels in same cluster is close compare to inter cluster relationship. In our proposed work we used k-means and EM segmentation algorithms for image segmentation, after segmenting the image features are extracted and classified using Feed forward neural network. Many research papers are describing the advancement of image processing for detecting the disease in plant leaf and variety of methodologies is described in the next section.

II. LITERATURE SURVEY

X.Agnes Kala Rani and R.Nagaraj[1] have proposed new method to find disease in tomato plant that is SOM based clustering for detecting bacterial spot disease in tomato field. In their research work, first images were aquired from field after that image proprocessing was done using four methods Gamma correction, Color space conversion, downsampling and Cutting for center parts is included after that images were segmented using SOM based clustering method result comarision was based on R2 comparision R2,R2adjusted, F-statistic and p-value comparison.

Saradhambal.G, Dhivya.R, Latha.S and R.Rajesh[2] have suggested a new method for Plant Disease Detection And Its Solution Using Image Classification.In their research work first they have collected images from real field after that noises were removed using pre-processing technique, after that image segmented using k-means segmentation algorithm and finally shape based features were extracted and finally diseases plants were identified using Otsu's classifier, after identifying the disease they have given solution for that.

Sagar Vetal and R.S. Khule [3] have proposed a new method for Tomato Plant Disease Detection using Image Processing. They did the following process to identify the diseases. First image acquisition then Image is smoothen using Kurtosis and skewness filters after that they have Perform the image segmentation using inverse difference method next feature extraction was performed, in this feature extraction they consider texture and color of an image to get the unique features from the image Finally they did classification using multi-class SVM algorithm. They got accuracy up to 93.75%.

Pradnya Ravindra Narvekar and S.N.Patil [5] have proposed a new method for Grape Leaf Disease Detection, Classification and Analysis by using Spatial Gray-level Dependence Matrices. First, the RGB images of leaves are converted into HSI color space In that HIS color space the H component is taken into account for further analysis after this they mask the green pixel after masking green pixel infected portion of the leaf was extracted then the infected region is segmented into a number of patches of equal size of 32*32 was taken the next step was to extract the useful segments the patches which are having more than fifty percent of the information were taken into account for the further analysis then they apply colour co-occurrence method to extract the features based on these extracted features they can easily classify black rot, downy mildew and powdery mildew diseases of grape leaf.

Ashish Miyatra and Sheetal Solanki [6] have suggested a new method for Disease and Nutrient Deficiency Detection in Cotton Plant. They mainly focus on Alternaria Leaf Spot disease and major nutrient deficiency detection, for segmentation they have used Statistical Region Merging algorithm and for Alternaria leaf spot disease detection they have used template matching algorithm and for detecting deficiency they have used color histogram algorithm.

Vijai Singh and A.K. Misra [10] have recommended a new method for Detection of plant leaf diseases using image segmentation and soft computing techniques. In their research work after getting image, green pixels were masked and removed after that genetic algorithm was used for segmentation then color co-occurrence method was used for feature extraction then classification was done using Minimum Distance Criterion and SVM classifier. The average accuracy of classification of their proposed algorithm was 97.6

III. PROPOSED SYSTEM

Tomato leaf images are collected to analysis whether the given leaf is healthy, show any deficiency or disease. First check the given leaf is healthy or not, if not identify deficiency shown or not, if it shown the deficiency what kind of deficiency and the result of that deficiency is found out. The proposed method starts with image acquisition then followed by image preprocessing after that image segmentation by k-means and Expectation and Maximization (EM) segmentation algorithms followed by feature extraction and image classification by support Vector Machine (SVM) classification algorithm. The following figure shows the proposed method.

Image Acquisition:

Tomato leaf images are collected from our Adhiyamaan College of Agriculture and Research, Hosur farm with digital camera with high resolution. The images are collected in a mixture way like healthy and deficiency showed. These images are used for further image processing techniques.

Preprocessing:

The collected images from image acquisition may contain images of different size, so first preprocessing step is to make all images in same size image for convenience. After resizing the image, noise in image must be removed, the noise may occur in image during capturing the image or transforming the image. Pixels intensity values changed from its true intensity value is called noise in the image. The process of removing noise from the noisy image using algorithm is called Noise removal. in this paper we use gaussian filter to remove noise from image. The resultant image after performing resize and noise removal operation is an enhanced image. These enhanced images are further used for image segmentation.

Retrieval Number: A1045109119/2019©BEIESP
DOI: 10.35940/ijeat.A1045.129219
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

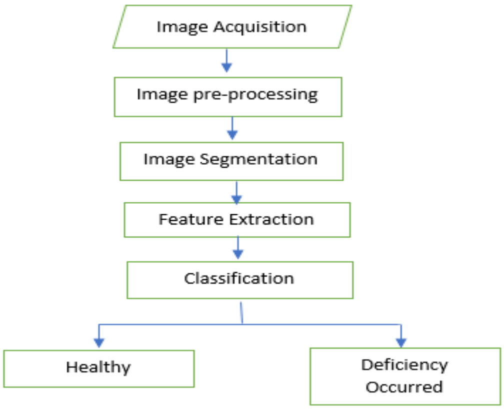


Fig 1. Block Diagram for the proposed work

K-means Segmentation:

Segmentation done by using k-mean clustering algorithm. One of the popular methods is k-means clustering for image segmentation. In k-means segmentation, the given image is partition into k number of segments with the relations that intra cluster pixels are closer than inter cluster pixels. K-means algorithm consists of two separate phases. To calculate the value for k and centroid for each segment is the first phase. In the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance. After the second phase the k-means algorithm recalculate the new centroid for each group based on the data available in that group. After calculating new centroid distance is calculated between each point and all the centroids and allots the points to the cluster which has minimum Euclidean distance. This process was repeated until no more changes between the clusters and now each cluster is considered as one segment and it is defined by its data and centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So, K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters. Let us consider an image with resolution of XxY and the image has to be cluster into k number of clusters. Let I be an Input image with set of pixels P (x, y) needs to be segmented and k and ck be the no of clusters and the cluster centers respectively. The algorithm for k-means clustering is following as:

1. Initialize number of cluster k and its center.
2. Calculate the Euclidean distance d, between the center and each pixel of an image using the relation given below.

$$d = \| p(x, y) - c_k \| \qquad (3)$$

3. Assign the pixels to the cluster where d is minimum.

4. After assigning all pixels to the cluster, recalculate new cluster center based on its elements using the relation given below.

$$c_k = \frac{1}{k} \sum_{y \in c_k} \sum_{x \in c_k} p(x, y) \qquad (4)$$

5. Repeat the process until no more changes in the clusters.

6. The cluster pixels are reshaped into image.

Expectation maximization:

Dimension of space, number of center and type of mixture models were specified initially and the value of variance kept unity. EM algorithm is used to efficiently calculate maximum likelihood even for hidden or missing data through its iterative procedure. When we reach either a maximum number of iterations to limit the time of calculation or a lower mistake then we can stop the algorithm. It is put easily in application because it leans on the calculation of the complete data. EM algorithm works in iterative way each iteration consists of two steps E-step and M-step. In the E-step missing data is replaced by estimated values. In M-step based on the assumption that data are known the likelihood function has been expanded. This implies that the EM algorithm ensures to increase the likelihood of data in every iteration. The following steps will clearly explain EM algorithm.

EM Algorithm:

1. Consider a set of starting parameters from the given a set of incomplete data
2. Expectation step (E – step): Estimate or guess the missing values from observed available values.
3. Maximization step (M – step): It is basically used to update the hypothesis. After Complete data generated from the expectation (E) step, it is used to update the parameters.
4. Step 2 and step 3 are repeated until convergence.

EM algorithm in detail.

- Initially, a set of initial values of the parameters are considered. A set of incomplete observed data is given to the system with the assumption that the observed data comes from a specific model.
- EM algorithm starts with considering initial values of the parameters. A set of incomplete data is also given with the assumption that it must be generated from E-step.
- Next step is Expectation step or *E-step*. In this step, based on observed data guess or estimate was made to estimate missing or incomplete data. It is basically used to update the variables.
- Next step is Maximization step or *M-step*. Based on the complete data generated from E-step, the values of the parameters are updated. It is basically used to update the hypothesis.
- In the last step, check is made whether the values are converging or not, if yes, then stop otherwise repeat *step-2* and *step-3* until the convergence occurs.

Feature Extraction:

Next step after the segmentation is interest features are extracted to identify the infected region. In this paper we used texture-based feature extraction technique for region description. Texture features like energy, entropy, contrast, standard deviation and variance are the most important features, the following texture features are used in our system.

Let M (i, j) be the co-occurrence matrix where i and j are the coefficients M (i, j), N is the element in the co-occurrence matrix at the coordinates i and j and N is the dimension of the co-occurrence matrix.

Energy

It mainly used to measure the amount of pixel pair repetitions based on energy we can easily measures the homogeneity of an image. The energy value will be large when the pixels are very similar, it is defined as

$$E = \sqrt{\sum_{i=0}^{N-1} \sum_{j=o}^{N-1} M^2(i, j)}$$

Entropy

Entropy (Ent) is mainly used to measure the randomness which is used to describe the texture of the input image. If all the elements of cooccurrence matrix are same then entropy value will be maximum. It is also defined as

$$Ent = \sum_{i=0}^{N-1} \sum_{j=o}^{N-1} M(i, j)(-\ln(M(i, j)))$$

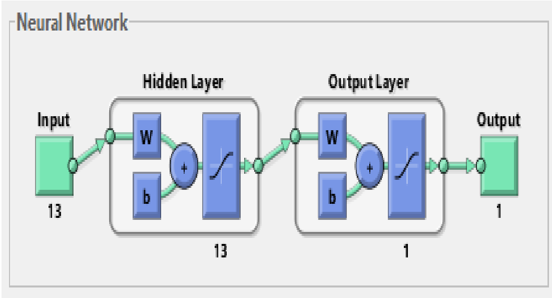
Contrast

The contrast (Con) is a measure of intensity of a pixel and its neighbor over the image. In the visual perception of the real world, contrast is determined by the difference in the color and brightness of the object and other objects within the same field of view.

$$Con = \sum_{i=0}^{N-1} \sum_{j=o}^{N-1} (i-j)^2 M(i, j)$$

Classification:

This is the final stage in deficiency identification. It is identifying a rule according to selected features and assigning each deficiency to any one the predetermined classes. In our proposed system we use Feed-forward back propagation network for classifying the deficiency. The most widely applied neural network algorithm in image classification remains the Feed-forward back propagation algorithm. Neural networks are composed of interconnected and interacting components called nodes or neurons. The feed-forward network has links that extend in only one direction. In this network there is one input layer attached with inputs similar to the number of features needed to predict the deficiency, some hidden layers that will forward the inputs through the output, no of hidden layers are depends upon the applications and the last layer is called output layer from this layer the result comes out as which deficiency occur. Each node in neural network is act as our biological neurons. Weight values are associated with each vector and node in the network, and these values constrain how input data are related to output data. The weight values will be adjusted during training phase of neural network to predict correctly. After getting trained, the neural network can be used for classification of new data.



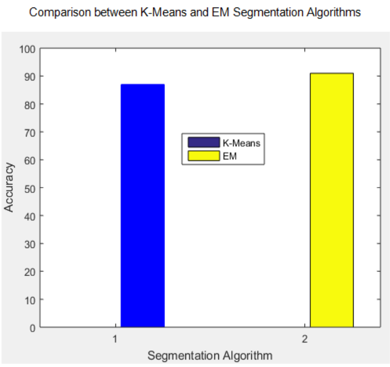
IV. RESULT AND DISCUSSION

Tomato leaves need 13 nutrients to grow among the 13 nutrients 3 nutrients are primary nutrients that are Nitrogen, phosphorus and potassium and 3 nutrients are secondary nutrients that are calcium, magnesium and sulfur and the remaining nutrients are micro nutrients that are Zinc, iron, boron, chloride, molybdenum, copper and manganese. Tomato plants need high amount of primary nutrients, little lower amount of secondary nutrients and very little amount of micro nutrients. Even though all these nutrients are important the primary and secondary nutrients are very important to produce good quality tomatoes. In our proposed system we consider primary and secondary nutrient deficiency alone not the micro nutrients. First images of each category were collected, in our research 10 images of each was considered for training and 3 images of each was consider for testing so totally 91 leaf images was considered that includes healthy leaf image also.

After performing preprocessing operation like resizing and noise removal operation, all training images are first segment using k-means segmentation algorithm and texture based features were selected like mean, standard deviation, variance, skewness and kurtosis after that images were classified using Feed forward neural network classification algorithm, after this testing was performed accuracy attained is 87%. The same process was repeated with EM segmentation algorithm and accuracy attained is 91%. Time consumed by MATLAB program using k-means segmentation algorithm is 132.172 seconds and using EM-Segmentation algorithm is 125.34 seconds, so based on execution time and accuracy the EM segmentation algorithm is more efficient than K-means segmentation algorithm.

V. CONCLUSION

In our proposed work two segmentation algorithms were used to find out deficiency in tomato leaves, when compare these two algorithms that is K-means and EM segmentation algorithm, the EM segmentation algorithm was more efficient that K-Means segmented algorithm. We tried to compare these algorithms with 91 images and we attain 87% accuracy in K-means and 91% accuracy in KM Segmentation. Further we will try to modify either k-means and EM algorithm to attain more accuracy as our future work.



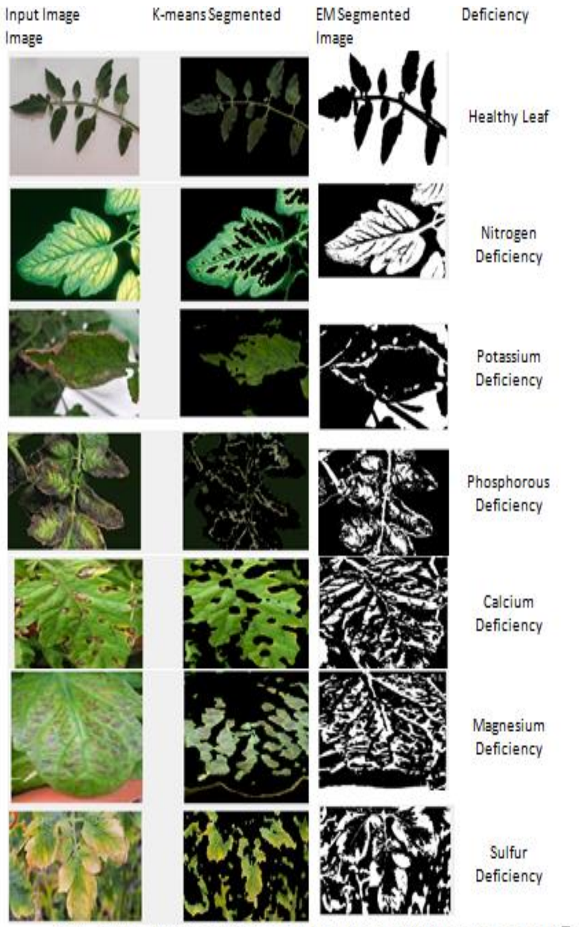


Fig 3. Nutrient deficiency detection using K-means and EM Segmentation algorithm

REFERENCES

1. X.Agnes Kala Rani and R.Nagaraj "SOM based clustering for detecting bacterial spot disease in tomato field" Indian Journal of Innovations and Developments Vol 5 (7), ISSN 2277-5382 July 2016
2. Saradhambal.G, Dhivya.R, Latha.S and R.Rajesh "Plant Disease Detection And Its Solution Using Image Classification" International Journal of Pure and Applied Mathematics Vol. 119 ISSN: 1314-3395 PP- 879-883
3. Sagar Vetal and R.S. Khule "Tomato Plant Disease Detection using Image Processing "International Journal of Advanced Research in Computer and Communication Engineering" Vol. 6, Issue 6, June 2017 pp-293-297
4. Sagar Karmarkar, Gauri Jadhav, Mayuri Parkhe, Aditya Kadam and Himali Patel "Tomato Disease Detection using Image Processing" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 7, Issue 12, December 2018 pp-4212-4220
5. Pradnya Ravindra Narvekar and S.N.Patil " Grape Leaf Disease Detection, Classification and Analysis by using Spatial Gray-level Dependence Matrices" International Journal of Advance Engineering and Research Development Vol. 1, Issue 9, September -2014 pp – 52-57
6. Ashish Miyatra and Sheetal Solanki "Disease and Nutrient Deficiency Detection in Cotton Plant" International Journal of Engineering Development and Research Vol. 2, Issue 2 pp-2801-2804 2014
7. Pawan P. Warne and Dr. S. R. Ganorkar " Detection of Diseases on Cotton Leaves Using K-Mean Clustering Method" Vol. 02 Issue: 04 pp-425-431 2015
8. Dheeb Al Bashish, Malik Braik and Sulieman Bani-Ahmad "Detection and Classification of Leaf Diseases using K-means-based Segmentation and Neural-networks-based classification" Information Technology Journal pp-267-275 2011
9. Prof. Sanjay B. Dhaygude and Mr.Nitin P.Kumbhar "Agricultural plant Leaf Disease Detection Using Image Processing" International Journal of Advanced Research in Electrical, Electronics and Instrumentation EngineeringVol. 2, Issue 1, January 2013 pp-599-602.
10. Vijai Singh and A.K. Misra "Detection of plant leaf diseases using image segmentation and soft computing techniques" Information Processing in Agriculture pp-41-49 2017

AUTHORS PROFILE

S.Sivagami, working as an Assistant Professor, department of Computer science, Adhiyamaan College of Agriculture and Research. She has completed her MCA and M.Phil., from Madurai Kamaraj University. She is Pursuing Ph.D in Periyar University Salem and her area of interest is Image Processing.

Dr. S.Mohanapriya working as an Associate Professor and head of department of Computer science, KSR College of Arts and Science for Women, Tiruchngode. She has completed her Ph.D at Anna University, Chennai and her area of interest is MANET and Image Processing