Smart Farming using K-means Clustering and SVM Classifier in Image Processing

Dr. SUJATHA S
Associate Professor, Dept. of TCE, CMR Institute of Technology, Bangalore, India

PREETI KUMARI
M.Tech Student, Dept. of TCE, CMR Institute of Technology, Bangalore, India

***Abstract*— Fruit being a major cultivable product of India is exported overseas so its quality is a major concern. The natural approach of quality assessment is done under human supervision, which is time consuming and quite expensive as it is manually done by experts. Therefore for quality assessment of fruit, a fully automated system is developed by using image processing technique. This will reduce human efforts, save time and increase the efficiency of fruit based industries where fruits are used in large quantities for making various fruit products. This approach can be used in import and export industry. In this paper, the main objective is to build the system which is fast, reliable, effective and error-free in order to detect the affected portion of fruit that helps agriculture for better quality and yield.**

***Keywords*—Image processing, K-means Clustering, SVM**

1) INTRODUCTION

Agriculture is the most important part of Indian economy. In India, agriculture contributes about sixteen percent (16%) of total GDP and ten percent (10%) of total exports. The classical approach for detecting the affected portion of fruit is based on the naked eye observation by the experts. Moreover consulting experts are expensive and time consuming in case of remote areas. Fruits have numerous uses in both fresh and artificial processed forms. The artificial processed form includes jam, pickles, juice.Fruit being a major cultivable product of India is exported overseas which results in increased revenue for the country. In order to get good quality of processed products the quality of fruits should also be good. The quality assessment of fruits in industries is done manually and it is main obstacle as it is time consuming and high labour cost.

Therefore it is very important to determine the quality of fruits for the purpose of its usage by an automatic sorting machine for various necessities in industries. To overcome this problem, image processing method may be used in an industry which has become a major source in recent years. Therefore, the main aim is to develop such a system which can detect the diseases in the fruits as soon as they produce their symptoms on the fruits such that proper management treatment can be applied.

The MATLAB image processing starts with acquiring the images from the digital high resolution camera or from the samples that is stored in the database. Affected and unaffected images of fruit are captured and stored for experiment. Then the images are applied for pre-processing in order to enhance the contrast of an image. Captured fruit images are segmented using k-means clustering method to form clusters. Features are extracted before applying K-means and SVM algorithm for training and classification. Finally diseases are recognized by this system and classified.

2) LITERATURE SURVEY

In paper [1] authors described image segmentation technique for plant disease detection. In image segmentation, K-means clustering algorithm is applied for separating foreground and background images. Clustering in segmentation is based on subtracting the clustered leaf images and intensity mapping for highlighting leaf area. K-means clustering is very effective and simple method for detection.

In paper [2] authors explained k-means clustering algorithm for infected leaf analysis. After applying k-means clustering techniques, the green pixel is masked and then by using Ostu's method, variable threshold value is obtained. For the feature extraction, color co-occurrence method (CCM) is used.

In paper [3] author uses image processing for detection of disease and the fruit grading. They have used artificial neural network for detection of disease. They have created two separate databases, one for the training of already stored disease images and other for the implementation of the query images. Back propagation is used for the weight adjustment of training databases. They consider three feature vectors, namely, color, textures and morphology.

In paper [4] author describes a novel defect of cashew nuts based on color and texture features with K-Nearest Neighbor algorithm. The Support Vector Machine (SVM) is used for background removal and color classification. Image processing is used to analyze the fruit's features; size, color, shape is determined based on the features.

In paper [5] author describes Layers separation technique is used for the training process which includes the training of these samples which separates the layers of RGB image into red, green, and blue layers and edge detection technique which detecting edges of the layered images. Spatial Gray-level Dependence Matrices are used for developing the color co-occurrence texture analysis method.

In paper [6] author first blur the image in order to reduce noise. Then the image is converted from RGB to HSV form, after this color thresholding is done. After thresholding foreground or background detection is performed.

Background detection leads to feature extractions of the leaf. Then k-means algorithm is applied which can help to clot the clusters.

In paper [7] addresses how the disease analysis is possible for the cotton leaf diseases detection, the analysis of the various diseases present on the cotton leaves can be effectively detected in the early stage before it will damage the whole plant. Artificial Neural Network as our classifier for comparison of cotton leaves. In this paper color is considered as feature.

3) PROPOSED WORK

The basic step of the proposed approach is shown in the fig 1.

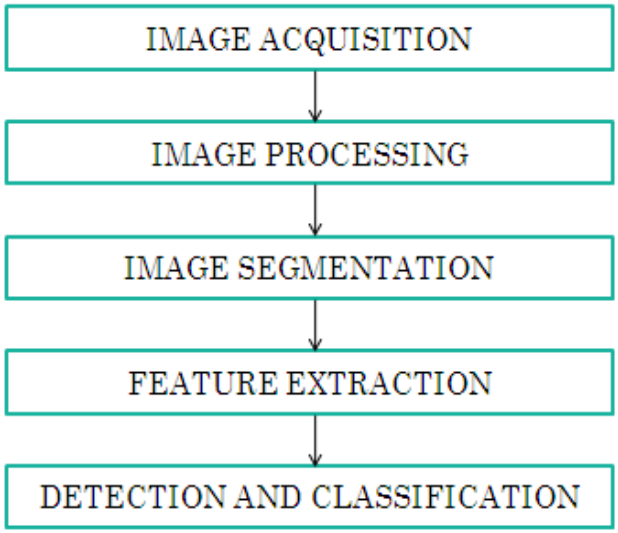

Fig. 1 Framework of the proposed system

Algorithm: Basic steps describing the proposed algorithm.

1. RGB image acquisition
2. Enhancing the contrast of an image
3. Apply K-means clustering method and select the region of interest (ROI).
4. Features are extracted by using GLCM Texture Statistics Computation
5. Configuring SVM Classifier for classification.
6. Accuracy and percentage of affected portion is calculated.

A. **Image Acquisition**

Image acquisition is the first method of digital image processing and it is described as capturing the image through digital camera or from the samples that are stored in database for further MATLAB operations.

B.**Image Pre-processing**

The main purpose of image pre-processing is to improve the quality of an image containing unwanted distortions or to enhance some image features for further processing. This method includes various techniques such as changing image size, filtering of noise, image conversion, enhancing image.

C**. Image Segmentation**

Segmentation means partitioning of image into various groups or clusters of same features or having some similarity. K-means clustering method partitions the images into clusters in such a way that atleast one part of cluster contain an image with major area of affected part.

Algorithm for the K-Means image segmentation:-
Step 1: Read input image.
*Step 2: Transform image from RGB to L*a*b* color space.*
*Step 3: Classify colors using K-Means clustering in 'a*b*' space.*
Step 4: Label each pixel in the image from the results of K-means.
Step 5: Generate images that segment the image by color.
Step 6.:Select the segment containing disease.

Fig. 2 Segmented region of interest

D.**Feature Extraction**

In feature extraction desired feature vectors such as color, texture, morphology and structure are extracted. Feature extraction is method for involving number of resources required to describe a large set of data accurately. Statistical texture features are obtained by Gray level co-occurrence matrix (GLCM) formula for texture analysis and texture features are calculated from statistical distribution of observed intensity combinations at the specified position relative to others. Numbers of gray levels are important in GLCM also statistics are categorized into order of first, second & higher for number of intensity points in each combination. Different statistical texture features of GLCM are energy, sum entropy, covariance, information measure of correlation, entropy, contrast and inverse difference and difference entropy.[3]

E. **Training & Classification**

Support vector machines (**SVMs)** are supervised learning models with associated learning algorithm that analyze data used for classification and regression analysis.[8]. Support vector machine is based on maximizing the minimum distance from the separating hyper plane to the nearest sample. Only binary classification is supported in basic SVM, but for multiclass classification case extension can be possible[8]. In these extensions, additional constraints and parameters are added to optimization problems for handling the separation of the different classes. SVM is a binary classifier that means the class labels can only take two values ±1. To get M-class classifiers, set of binary classifiers are constructed in this way $f^1 f^2 \ldots\ldots f^M$ and each is trained for separating one class from the rest.

$$\underset{j=1\ldots M}{\operatorname{argmax}}\, g^j(x) \text{ , where } g^j(x) = \sum_{i=1}^{m} y_i \alpha_i^j k(x, x_i) + b^j$$

The $g^j(x)$ function returns the signed real value that can be interpreted as distance from separation of hyper plane to point x.. The larger the value the more confident one is that the point x belong to the positive class. Hence, assign point x to the class whose confidence value is largest for this point.

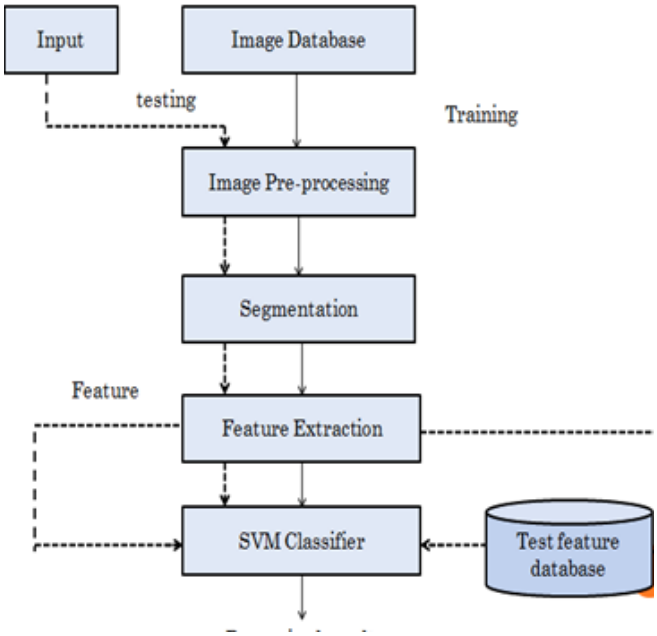

Fig. 3 Principle of SVM

4) RESULTS

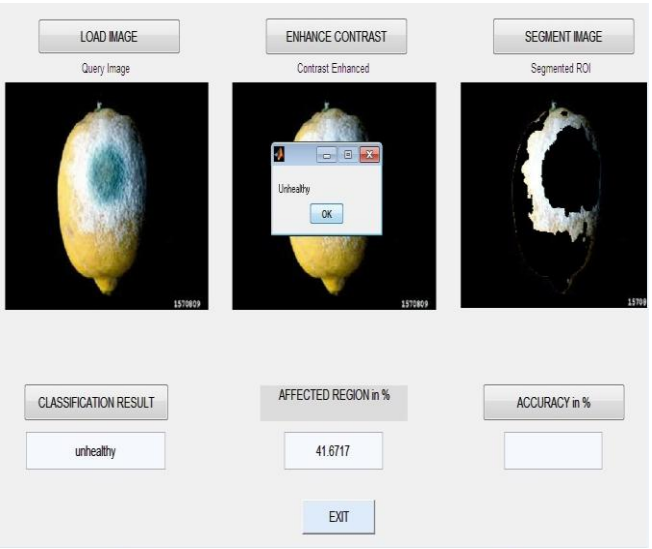

Fig 4 System shows classification result

5) CONCLUSION

An image processing based solution is proposed and evaluated for the detection and classification of affected portion of fruit. This paper provides an efficient and accurate system for detection and classification of fruit images which is affected. The proposed approach uses K-Means clustering and SVM classifier, as it provides high accuracy when compared to other techniques and consumes very less time for entire processing.

6) REFERENCES

1) Sachin D. Khirade, A. B. Patil, "Plant Disease Detection Using Image Processing" IEEE International Conference on Computing Communication Control and Automation, pp. 978-1-4799-6892-3/15, 2015.

2) Mrunalini R. Badnakhe, Prashant R. Deshmukh, "Infected Leaf Analysis and Comparison by Otsu Threshold and k-Means Clustering", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 3, March 2012.

3) Monika Jhuria, Rushikesh borse, Ashwani Kumar "Image Processing for Smart Farming: Detection of Disease and Fruit Grading" Proceeding of the IEEE Second International Conference on Image Information Processing, pp. 978-1-4673-6101, 2013.

4) Miss. Vrushali Nagpure," Grading of Cashew Nuts on the Bases of Texture, Color and Size", International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 4 Issue: 4 171 - 173 .

5) Smita Naikwadi, Niket Amoda" Advances in Image Processing for Detection of Plant Diseases" , International Journal of Application or Innovation in Engineering & Management , Volume 2, Issue 11, November 2013.

6) Vaijinath B. Batule," Leaf Disease Detection using Image Processing and Support Vector Machine(SVM)", Journal for Research| Volume 02| Issue 02 | April 2016 ISSN: 2395-7549.

7) Malvika Ranjan," Detection And Classification Of Leaf Disease Using Artificial Neural Network", International Journal of Technical Research and Applications e-ISSN: 2320-8163, www.ijtra.com Volume 3, Issue 3 (May-June 2015), PP. 331-333.

8) Pankaj Bhowmik ,"Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine, 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE).

9) Ms. Seema Banot, Dr. P.M. Mahajan," A Fruit Detecting And Grading System Based On Image Processing-review", International Journal Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering Vol. 4, Issue 1, January 2016 .

10) Sowmya GM, "Disease Detection In Pomegranate Leaf Using Image Processing Technique", International Journal of Science, Engineering and Technology Research (IJSETR) Volume 6, Issue 3, March 2017, ISSN: 2278 -7798.

Sujatha.S received the B.E. Degree in Telecommunication Engineering from MVJ College of Engineering, Bangalore, India, in 2001, the M.Tech. Degree in VLSI and embedded system design from the BMS College of Engineering Bangalore, India, in 2010, and Ph.D. degree in Electronics and Communication Engineering from, Pondicherry Engineering College, Pondicherry, India. She is working as an Associate Professor in the Department of Telecommunication CMR institute of Technology since 2007 till date. Her research interests include Wireless Communication, Optical Communication, Optical Communication Network, Microcontroller, and MSP430. She has presented and published several papers in National and International Journal.

Preeti Kumari received the B.E Degree in Electronics and Communication Engineering from East Point College of Engineering, Bangalore, India, in 2016, pursuing M.Tech Degree in Digital Communication from CMR Institute of Technology Bangalore, India. Her research interests include image processing, networks.

.