ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 4, April 2016

DETECTION AND CLASSIFICATION OF FRUIT DISEASE BY USING COMPLETE LOCAL BINARY PATTERN

Mrs.R.Vidya
Assistant Professor, ECE.
S.A Engineering College
Chennai, India

Ms.S.Sabitha
Student, ECE
S.A Engineering College
Chennai, India

Ms.A.Soundarya
Student, ECE
S.A Engineering College
Chennai, India

Abstract **Diseases in fruit causes devastating problem in economic losses and production in agricultural industry worldwide. In this paper, a solution for the detection and classification of fruit diseases is proposed and experimentally validated. The proposed system involve K-means clustering which is used for image segmentation, some state of the art features are extracted from the segmented image and finally images are classified into one of the classes by using a Multi-class Support Vector Machine. Our experimental result express that the proposed solution can significantly support accurate detection and automatic classification of fruit diseases. The classification accuracy for the proposed solution is achieved up to 93%**

Keywords— K-means clusteing, Local Binary Pattern, Multi-class Support Vector Machine, Texture classification

I.INTRODUCTION

Images may be two-dimensional, such as a photograph, screen display and as well as a three-dimensional, such as a cameras, mirrors, lenses, telescopes, microscopes, etc. and natural objects and phenomena, such as the human eye or water. A volatile image is one that exists only for a short period of time. This may be a reflection of an object by a mirror, a projection of a camera obscura or a scene displayed on a cathode ray rube. A fixed image, also called a hard copy, is one that has been recorded on a material object such as paper or textile by photograph or any other digital process. A mental image exist in an individual remember or imagines. The image object need not be real it is an abstract concept such as graph. A single static image is distinguished from a kinetic image. A photograph is taken on the set of a movie or television program during production used for promotional purposes.

A digital image is also represented in two-dimensional image in numeric representation. It is also termed as raster image or bit mapped image. Photo is an image created by light. The optical and analog image processing is possible in digital image. Image processing also refers to digital image. The ever growing importance of visualization is due to the image gain much broader scopes.

In short period of time the volatile image will display. Image acquisition, image enhancement, image restoration, image segmentation are the steps in digital image processing. The Image height and width are definitely counted in pixels.

II. RELATED WORK

Author [1] Diseases of apple fruit appearing at harvest can cause significant losses in yield and quality. To know what control measures to take next year to prevent similar losses. It is important to recognize what is being observed. In some case the growers will need to cut the fruit open to identify the problem. Many of the diseases of apple fruit also attack other parts of the tree causing diseases of leaves, twigs, and branches. These tissues are sources of inoculums for the fruit diseases, some of the apple fruit diseases also attack wild plants such as brambles and woodland species growing nearby. These native plants are also a source of inoculums for diseases. It describes the most important fungus-caused diseases of apple fruits in Kentucky. For more complete descriptions, with helpful color pictures, growers are encouraged to refer to the compendium of Apple and pear diseases.

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 4, April 2016

Author [2] Dubey.S The proposed approach used k-means clustering technique for segmenting defects with three or four clusters. We have used defected apples for the experimental observations and evaluated the introduced method considering apples as a case study. The proposed approach is able to accurately segment the defected area of fruit presented in the image. K-means based defect segmentation approach is also segment defected area with the stem and calyx of the fruits. This work presents a novel defect segmentation of fruit based on color features with K-means clustering unsupervised algorithm. We used color images of fruits for defect segmentation. Defect segmentation is carried out into two stages. At first, the pixels are clustered based on their color and spatial features, where the clustering process is accomplished. Then the clustered blocks are merged to a specific number of regions. Using this two step procedure, it is possible to increase the computational efficiency avoiding feature extraction for every pixel in the image of fruits. This approach thus provides a feasible robust solution for defect segmentation of fruits.

III.PROPOSED SYSTEM

In addition to the existing system the image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. The accuracy will improve.

IV.CLUSTERING METHODS

The K-means algorithm is an iterative technique that is used to partition an image into K clusters The basic algorithm is:

- Pick K cluster centers, either randomly or based on some heuristic
- Assign each pixel in the image to the cluster that minimizes the distance between the pixel and the cluster center.
- Re-compute the cluster centers by averaging all of the pixels in the cluster
- Repeat steps 2 and 3 until convergence is attained

In this case, distance is the squared or absolute difference between pixel and a cluster center. The difference is typically based on pixel color, intensity, texture, and location, or a weighted combination of these factors. K can be selected manually, randomly, or by a heuristic.

This algorithm is guaranteed to converge, but it may not return the optimal solution. The quality of the solution depends on the initial set of clusters and the value of K.

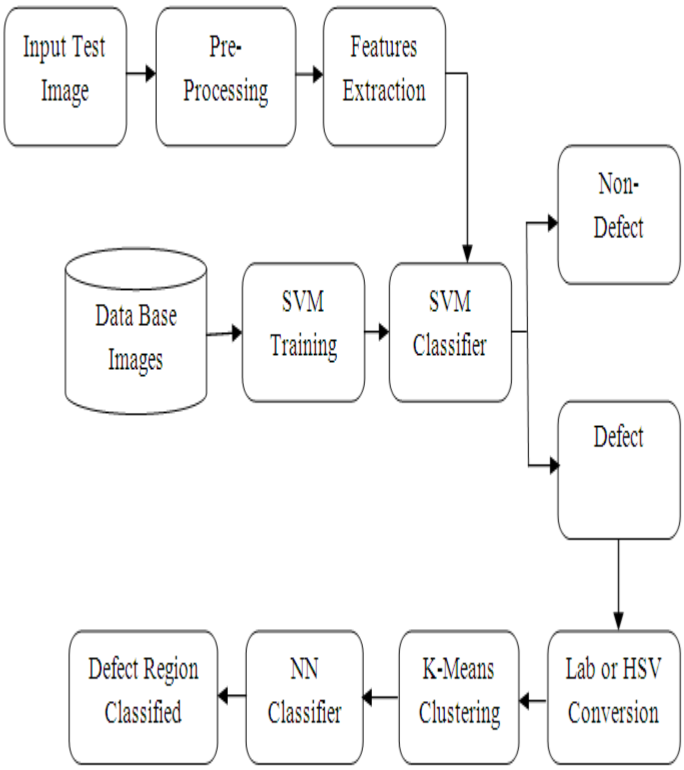
The k-means clustering was invented in 1956. The most common form of the algorithm uses an iterative refinement heuristic known as LIoyd's algorithm. LIody's algorithm starts by partitioning the input points into k initial sets, either at random or using some heuristic data. It then calculates the mean point, or centroid. Then the centroid is recalculated for the new clusters, and algorithm repeated by alternate application of these two steps until convergence, which is obtained when the points no longer switch clusters. Lloyd's algorithm is a heuristic for solving the K-means problem, as with certain combination of starting point and centroid, Lloyd's algorithms can in fact coverage to the wrong answer.

Other variations exist, but Lloyd's algorithm has remained popular, because it converges extremely quickly in practice. In terms of performance the algorithm is not guaranteed to return a global optimum. The quality of the final solution depends largely on the initial set of clusters, and many in practice be much poor than the global optimum. Since the algorithm is extremely fast, a common method is to run the algorithm several times and return the best clustering found.

A drawback of the K-means algorithm is that the number of clusters K is an input parameter. An inappropriate choice of K may yield poor result. The algorithm also assumes that the variance is an appropriate measure of cluster scatter

V. PREPROCESSING

Preprocessing is an important and diverse set of image preparation programs that act to offset problems with the band data and recalculate DN values that minimize these problems.



Among the programs that optimize these values are atmospheric correction sun illumination geometry; surface-induced geometric distortions; spacecraft velocity and attitude variations ; effects of Earth rotation, elevation, curvature, abnormalities of instrument performance ; loss of specific scan lines, and others.

Once performed on the raw data, these adjustments require appropriate radiometric and geometric corrections.

Re sampling is one approach commonly used to produce better estimates of the DN values for individual pixels However, the radiometric values of the displaced pixels no longer represent the real world values that would be obtained if this new pixel array could be resented by the scanner.

The particular mixture of surface objects or materials in the original pixel has changed somewhat depending on pixel size, extent on continuation of these features in neighboring pixels.

In simple words the corrections have led to a pixel that at the time of sampling covered. In the nearest neighboring technique, the transformed pixel takes the value of the closest pixel I the pre shifted array. In the biller interpolation approach, the average of the DNs for the four pixels surrounding the transformed output pixel is used.

VI. FILTER

Preprocessing produces fall into three broad categories. They are image restoration, image enhancement and classification and information extraction. Under information extraction rationing and principle component analysis has elements of enhancement but lead to image that can be interpreted directly for recognition and identification of classes and features. Also include the third category but treated outside is change detection. The radiance, such as reflectance's and remittance's which vary through a continuous range of values are digitized on board the spacecraft after initially being measured by the sensor in use. Ground instrument data can also be digitized at the time of collection or imaginary obtained by conventional photography is capable of digitization. A DN is simply one of the set of numbers based on powers of 2.

The range of radiances which instrument-wise , can be for example , recorded as varying voltages if the sensor signal is one which is say the conversion of photos counted at a specific wavelength or wavelength intervals. The lower and upper limits of the sensor's response capability form he end numbers of the DN range selected. The voltages are divided into equal whole number units based on the digitizing range selected. Thus a land sat TM band can have its voltage value the maximum and minimum that can be measured subdivided into 2^8 or 256 equal units. These are arbitrarily set as 0 for the lowest value, so the range is then 0 to 255. It is mainly used to suppress either the high frequency in image.

There are two domains in filter they are frequency domain and spatial domain. In these domains the images are filtered. The first involves transforming the image into the frequency domain, multiplying with the frequency filter function and re-transforming the result into the spatial domain. The filter function is shaped so as to attenuate some frequencies and enhance others. For example, a simple low pass function is 1 for frequencies smaller than the cut-off frequency and 0 for all others. The corresponding process in the spatial domain is to convolve the input image. The mathematical operation is identical to the multiplication in the frequency space, but results of the digital implementations vary, since approximate filter function with a discrete and finite kernel. It has some specific application, where the size and the form of the kernel determine the characteristics of the operation. The kernel for two examples mean and the laplacian operator.

Filter operation is called feature extraction. To extract the features from the image Color Histogram features, Color Coherence vector features and Local Binary Pattern features are extracted from the image. The color channels of the images are separated and histogram is applied to each color channels. The values are saved as feature. The Color Coherence Vector is calculated for the image and the values are stored as features. Then finally LBP features are obtained by the comparison of the pixels with the neighboring pixels and the values are saved as features.

VII. SUPPOR VECTOR MACHINE

Improving classifier effectiveness has been an area of intensive machine-learning research over the last two decades, and this work has led to a new generation of state-of-art classifier, such as support vector machines, boosted decision trees, regularized logistic regression, neural network, and random forests. Many of these methods, including support vector machines, the main topic of this chapter, have been applied with success to information retrieval problems, particularly text classification. An SVM is a kind of large-margin classifier. It is a vector space based machine learning method where the goal is to find a decision boundary between two classes that is maximally far from any point in the training data. We will initially motivate and develop SVMs for the case of two-class data sets that are separable by a linear classifier and then extend the model in to non-separable data, multi-class problems, and nonlinear models, and also present some additional discussion of SVM performance.

The representation of SVM model is of the example of the separate categories is divided by a clear gap that is as wide as possible. Linear classification of SVMs can efficiently perform a non-linear classification is called the kernel trick. When data are not labeled, a supervised learning is not possible and an unsupervised learning is required, that would find natural clustering of the data to groups, and map new data to these formed groups. The clustering algorithm which provides an improvement to the support vector machine is called support vector clustering and is often used in industrial applications either when only some data is labeled as a preprocessing for a classification passes. Support vector machine constructs a hyper plane or set of hyper planes in a high or in finite dimensional space, which can be used for classification, regression, or other tasks.

To keep the computational load reasonable the mapping used by SVM schemes are designed to ensure that dot product may be computed easily in terms of the variables supervised learning with non knowledge based classifier will be used for image classification. The SVM is used here to act as a classifier with radial basis function for network activation function. The training samples features with assigned target vectors are fed into created SVM model for supervised training to get machine parameters such as node biases and weighting factors. Finally, test image features are simulating with trained machine learning parameter to make decision of fruit is normal or abnormal. SVM can be used to solve various real world problems. SVMs are helpful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings. SVM are useful in medical science to classify proteins with up to 90% of the compounds classified correctly. SVM achieve significantly higher search accuracy than traditional query refinement schemes.

VIII. CONCLUSION AND FUTURE WORKS

In this paper the detection of fruit disease is proposed and evaluated by using complete local binary pattern. This approach involve three methods K-means clustering technique which is used for image segmentation, in which image features are extracted from the segmented image, finally the images are trained and classified by using multi-class support vector machine. Experimental results indicate that the proposed solution can significantly support accurate detection and automatic classification. The proposed system found that the normal fruits are easily distinguishable with the diseased fruit and LBP feature shows more accurate result for the identification of fruit disease and achieved more than 93% classification accuracy. Further work includes consideration of fusion of more than one feature to improve the output of the proposed method. Histogram chain code is better result for the classification of fruit disease and achieved more than 98% classification accuracy. Further work includes consideration of fusion of more than one feature to improve the output of the proposed methods. The accuracy of the proposed approach is defined as,

$$\text{Accuracy}(\%) = \frac{\text{Total num of images correctly classified}*100}{\text{Total number used for testing}}$$

Images are tested in multi-class support vector machine using histogram of chain code and density of pixel features: choose the linear kernel, polynomial kernel, quadratic kernel amd radian basis kernel.

For future study, further different neural network architectures, SVM, fuzzy based classifier can be used for classification. We can extend this work to classify fungal disease symptoms affected on commercial crops, cereals like viral, bacteria affected on agriculture/horticulture produce.

IX. REFERENCE

[1] Roberts, M. J., Schimmelpfennig, D., Ashley, E., Livingston, M., Ash, M., & Vasavada, U. (2006). *The Value of Plant Disease Early- Warning Systems* (No. 18).Economic Research Service, United States Department of Agriculture.

[2] Hartman, J. (2010, April). *Apple Fruit Diseases Appearing at Harvest.* Plant Pathology Fact Sheet, College of Agriculture, University of Kentucky.

[3] Li, Q., Wang, M., & Gu, W. (2002, November). Computer Vision Based System for Apple Surface Defect Detection. *Computers and Electronics in Agriculture*, 36, 215-223.

[4] Mehl, P. M., Chao, K., Kim, M., & Chen, Y. R. (2002). Detection of Defects on Selected Apple Cultivars using Hyperspectral and Multispectral Image Analysis. *Applied Engineering in Agriculture*, 18, 219-226.

[5] Kim, M. S., Lefcourt, A. M., Chen, Y. R., & Tao, Y. (2005). Automated Detection of Fecal Contamination of Apples Based on Multispectral Fluorescence Image Fusion. *Journal of food engineering*, 71, 85-91.

[6] Kleynen, O., Leemans, V., & Destain, M. F. (2005). Development of a Multi-Spectral Vision System for the Detection of Defects on Apples. *Journal of Food Engineering*, 69, 41-49.

[7] Leemans, V., Magein, H., & Destain, M. F. (1999). Defect Segmentation on _Jonagold‘ Apples using Color Vision and a Bayesian Classification Method. *Computers and Electronics in Agriculture*, 23, 43-53.

[8] Leemans, V., Magein, H., & Destain, M. F. (1998). Defect Segmentation on _Golden Delicious‘ Apples by using Color Machine Vision. *Computers and Electronics in Agricultur*e, 20, 117-130

[9] Bravo, C., Moshou, D., Oberti, R., West, J., McCartney, A., Bodria, L., & Ramon, H. (2004). Foliar Disease Detection in the Field using Optical Sensor Fusion. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development*, 6, 1-14.

[10] Moshou, D., Bravo, C., Wahlen, S., West, J., McCartney, A., De, J., Baerdemaeker, J. D., & Ramon, H. (2006). Simultaneous Identification of Plant Stresses and Diseases in Arable Crops using Proximal Optical Sensing and Self-Organising Maps. *Precision Agriculture*, 7(3), 149-164

[11] Spinelli, F., Noferini, M., & Costa, G. (2006). Near Infrared Spectroscopy (NIRs): Perspective of Fire Blight Detection in Asymptomatic Plant Material. In *Proceedings of the 10th International Workshop on Fire Blight*, *Acta Horticulturae* (pp. 87-90).

[12] Yang, C. M., Cheng, C. H., & Chen, R. K. (2007). Changes in Spectral Characteristics of Rice Canopy Infested with Brown Planthopper and Leaffolder. *Crop Science*, 47, 329-335.

[13] Chen, B., Wang, K., Li, S., Wang, J., Bai, J., Xiao, C., & Lai, J. (2008). Spectrum Characteristics of Cotton Canopy Infected with Verticillium Wilt and Inversion of Severity Level. In *Proceedings of the Computer and Computing Technologies in Agriculture* (Vol 2, pp. 1169-1180).

[14] Choi, Y. H., Tapias, E. C., Kim, H. K., Lefeber, A. W. M., Erkelens, C., Verhoeven, J. T. J., Brzin, J., Zel, J., (2004). Metabolic Discrimination of Catharanthus Roseus Leaves Infected by Phytoplasma using 1H-NMR Spectroscopy and Multivariate Data Analysis. *Plant Physiology*, 135, 2398-2410.

[15] Hahn, F. (2009). Actual Pathogen Detection: Sensors and Algorithms—A Review. *Algorithms*, 2(1), 301-338.

[16] Sankarana, S., Mishraa, A., Ehsania, R., & Davisb, C. (2010). A Review of Advanced Techniques for Detecting Plant Diseases. *Computers and Electronics in Agriculture*, 72, 1-13.

[17] Moshou, D., Bravo, C., Oberti, R., West, J., Bodria, L., McCartney, A., & Ramon, H. (2005). Plant Disease Detection Based on Data Fusion of Hyper-Spectral and Multi-Spectral Fluorescence Imaging using Kohonen Maps. *Real-Time Imaging*, 11(2), 75-83.

[18] Shafri, H. Z. M., & Hamdan, N. (2009). Hyperspectral Imagery for Mapping Disease Infection in Oil Palm Plantation using Vegetation Indices and Red Edge Techniques. *American Journal of Applied Sciences*, 6(6), 1031-1035.

[19 Qin, J., Burks, F., Ritenour, M. A., & Bonn, W. G. (2009). Detection of Citrus Canker using Hyper-Spectral Reflectance Imaging with Spectral Information Divergence. *Journal of Food Engineering*, 93(2), 183-191.

[20] Purcell, D. E., O'Shea, M.G., Johnson, R. A., & Kokot, S. (2009). Near-Infrared Spectroscopy for the Prediction of Disease Rating for Fiji Leaf Gall in Sugarcane Clones. *Applied Spectroscopy*, 63(4), 450-457.

[20] Marcassa, L. G., Gasparoto, M. C. G., Belasque, J., Lins, E. C., Nunes, F. D., & Bagnato, V. S. (2006). Fluorescence Spectroscopy Applied to Orange Trees. *Laser Physics*, 16(5), 884-888.

[21] Belasque, L., Gasparoto, M. C. G., & Marcassa, L. G. (2008). Detection of Mechanical and Disease Stresses in Citrus Plants by Fluorescence Spectroscopy. *Applied Optics*, 7(11), 1922-1926.

[22] Guo, Z., Zhang, L., & Zhang, D. (2010). A Completed Modeling of Local Binary Pattern Operator for Texture Classification. *IEEE Transactions on Image Processing (TIP)*, 19(6), 1657-1663.

[23] Ojala, T., Pietikäinen, M., & Mäenpää, T. T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Pattern. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 24(7), 971-987.

[24] Marszalek, M., & Schmid, C. (2006). Spatial Weighting for Bag-of-Features. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 2118-2125).

[25 Grauman, K., & Darrel, T. (2005). Efficient Image Matching with Distributions of Local Invariant Features. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 627-634).

[26] Rocha, A., Hauagge, C., Wainer, J., & Siome, D. (2010). Automatic Fruit and Vegetable Classification from Images. *Computers and Electronics in Agriculture*, 70, 96-104.

[27] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, *1st edition*, Springer. Dubey, S. R. (2012). Automatic Recognition of Fruits and Vegetables and Detection of Fruit Diseases. *Master's theses*, GLA University Mathura, India. Dubey, S. R., & Jalal, A. S. (2012a). Robust Approach for Fruit and Vegetable Classification. *Procedia Engineering*, *38*, 3449 – 3453.

[28] Dubey, S. R., & Jalal, A. S. (2012b). Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns In *Proceedings of the 3rd International Conference on Computer and Communication Technology* (pp. 346-351), MNNIT Allahabad, India.

[29] Dubey, S. R., & Jalal, A. S. (2013). Species and Variety Detection of Fruits and Vegetables from Images. *International Journal of Applied Pattern Recognition*, *1*(1), 108 – 126.

[30] Dubey, S. R., Dixit, P., Singh, N., & Gupta, J. P. (2013). Infected fruit part detection using K-means clustering segmentation technique. *International Journal of Artificial Intelligence and Interactive Multimedia*, *2*(2).

