A Survey on Apple Fruit Diseases Detection and Classification

Bhavini J. Samajpati
M.E Student of Computer Department
Sigma Institute of Engineering, Vadodara

Sheshang D. Degadwala
Assistant Professor of Computer Department
Sigma Institute of Engineering, Vadodara

ABSTRACT
Images are the essential source of information and data in agribusiness science. There is a mesh criticalness of farming in India. The nature of organic product assumes a key part in agro based applications. Early detection of infection and crop health can provide the control of fruit diseases through legitimate administration approaches. Human administrators inspect the organic product by outwardly which is monotonous and tedious procedure. So machine vision and image processing procedures are utilized. This paper surveys the methodologies utilized for apple fruit diseases detection, Segmentation of infected apple fruit part and classification of diseases by using image processing. Likewise states summary of various color techniques, various texture techniques, various segmentation techniques and various classifiers all with their benefits and negative marks.

Keywords
Color features, texture features, classifier, segmentation techniques

1. INTRODUCTION
India has a second rank in the production of fruit[2]. Agriculture is a significant sector of Indian economy because it contributes 17 percent of the total Gross Domestic Product(GDP) of India and provides the employment to over 60 percent population. Consistently huge amounts of apples are delivered and reaped all through world. To deliver great quality of products it is important to discover infections in a fruit. The traditional methodology for diseases detection and identification of fruit depends on the exposed eye perception by the specialists yet in some creating nations, counseling specialists are costly and tedious because of the far off areas of their availability. Automatic identification of fruit diseases is crucial to consequently identify the manifestations of diseases as right on time as they show up in developing natural product. Apple fruit diseases can cause major losses in yield and quality appeared in harvesting[3].Fig 1 shows some common diseases of apple fruits are apple scab, apple rot and apple blotch[1].Apple scab are gray or brown corky spots. Apple rot is having slightly sunken, circular brown or black spots which may covered by a red halo. Apple blotch diseases is a fungal diseases which is appears on the surface of the apple fruit as dark, irregular or lobed edges[1].

Visual investigation of apples is as of now computerized in the business by machine vision concerning size and color.
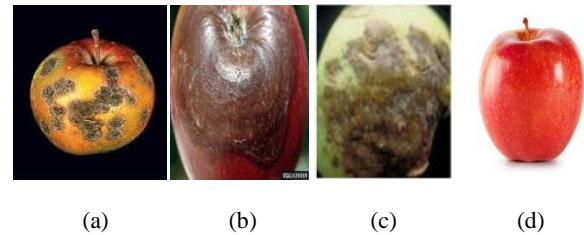

(a)   (b)   (c)   (d)

Fig 1 : Three common apple fruit diseases: (a) apple scab (b) apple rot(c) apple blotch (d) normal apple.

Detection of disease is still a testing undertaking because of common variability of skin color in diverse types of apples, high variability of defect types, and vicinity of stem/calyx.

In this paper it is reviewed that the detection of three common apple fruit diseases and classification of detected diseases of apple fruit.

2. DISEASES DETECTION AND CLASSIFICATION
Features are extracted from the fruit because based on its features able to classify fruit diseases. Extract various features from fruit like color and texture features.

In apple fruit diseases detection process identifying which type of diseases available in the apple fruit. The diseases detection of apple fruit is basically having three steps. Fig 2 shows the basic approach for detecting the diseases of apple fruit.

The first and the basic step is to prepare the training and testing data set of the apple fruit images. Secondly segment the infected apple fruit part from the testing image by using segmentation techniques. Third step is feature extraction. In this step some color and texture features are extracted to identify the apple diseases. Finally the detected diseases are classified by using classifiers.

Shivram Dubey et al(and other)(2012)[1].In this paper three apple diseases have been taken apple scab, apple rot and apple blotch[1]. The first step is image segmentation by K-mean clustering. In second step features are extracted. These features are extracted from the segmented image and that features are global color histogram(GCH),color coherence vector(CCV),local binary pattern(LBP) and complete local binary pattern(CLBP).In which complete local binary pattern give 93% correct classification accuracy than other features because it calculates magnitude, sign and centre value of pixel.
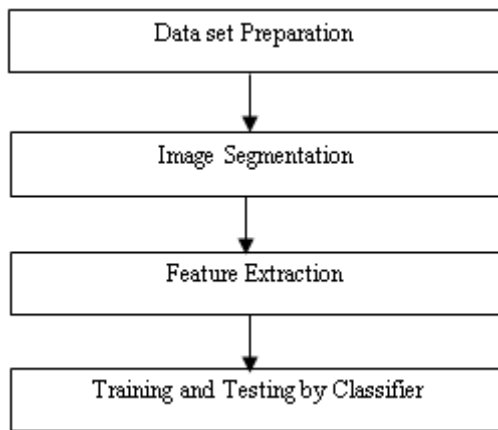
Fig 2: Steps to detect diseases and classification of diseases from apple fruit.

Shivaram Dubey et al(and others)(2013)[4].This paper displays a novel defect segmentation of fruit based on color component with K-means clustering unsupervised algorithm. They utilized color image of fruit for defect segmentation. Defect segmentation is done into two stages. At to start with, the pixels are clustered based on their color and spatial features, where the clustering procedure is accomplished. At that point the clustered blocks are converged to a particular number of region. Utilizing this two stage system, it is possible to increase the computational efficiency avoiding feature extraction for each pixel in the image of fruit. Likewise expresses a survey of the different segmentation techniques. The fundamental point of this methodology is to segment colors automatically utilizing the K-means clustering techniques and L*a*b* color space.

R.Sivamoorthi et al(and other)(2015)[5].In this paper they have proposed a methodology which utilizes neural system to recognize the diseases in the apple fruit. The features are extracted from the images utilizing Color Invariant, Color histogram and Local Binary pattern algorithm. The features are extracted for the input image and all of the images in the database. The extracted feature values are gone into the neural network classifier. The neural network classifies the image into ordinary or defected. If the defected part is recognized means the defected region is segmented. The color image is changed over L*a*b* Color Space. At that point K-means clustering is applied to classify the images. At that point at last the defected part is detected by recognizing the defected cluster. The extracted test image features and the train image features are then gone into the Multi class SVM classifier so as to discover the kind of defects in the cluster.

Priya P.Donny et al(and other)(2015)[6]. This article depicts the feature extraction methods for crop and fruit infections based on computer image processing in detail. The primary motivation behind this paper is to give an interface to digitally illiterate clients, particularly farmers to efficiently and effectively recover data through web. Likewise, to enable the farmers to recognize the diseases in their yield, its reasons and symptoms utilizing image processing without classical approach and identify the diseases. Firstly by color transformation structure RGB is changed over into HSV space because HSV is a decent color descriptor. Masking and removing of green pixels with pre-computed threshold level. At that point in the following step segmentation is performed utilizing k-means clustering technique .These sections are utilized for texture analysis by color color co-occurrence matrix. Finally texture parameters are compared to texture

parameters of normal fruit image. The four unique types of extraction strategies for effective investigation are concentrated to have efficient detection and classification.

Jagadeesh. D. Pujari et al (and other)(2013)[7]. In this paper, they have displayed a reduced feature set based methodology for recognition and classification image of fruits into normal and defected. The RGB (Red Green Blue) color features are reduced from 18 to 2 and GLCM (Gray-level Co-occurrence Matrix) texture features are reduced from 30 to 2. The reduced features set of comprises of 4 features namely, green mean, saturation mean, red GLCM sum mean and green GLCM sum mean. A feedback from classifier, performance is utilized in reducing the features. The average accuracy of 89.15% for normal type and 88.58% for affected type is obtained utilizing 2 color features. The average accuracy of 93.15% for normal type and 89.50% for affected type is obtained utilizing 2 texture features. The average accuracy have expanded to 96.85% for normal type and 93.89% for defected type when the reduced color features and texture features are joined. The work discovers application in developing a machine vision system in agriculture. A BPNN classifier is suitable for this work.

Monika Jhuria et al (and other)(2013)[8]. In this paper image processing is utilized as tool to monitor the execution on fruits during cultivating, right from plantation to reaping. For this reason artificial neural network is utilized. Three defects of grapes fruit and two of apple fruit have been picked. The framework utilizes two image databases, one for training of already stored defected images the other for implementation of query images. Back propagation method is utilized for weight adjustment of training images database. The images are categorized and mapped to their respective diseases categories on reasons of three features vectors namely to be specific, color, texture and morphology. From these feature vector morphology gives 90% right outcome and it is more than other two feature vectors. This paper exhibits effective algorithm for spread of diseases and mango counting. The implementation of neural networks has been done using MATLAB.

Wang Xingyuan et al(and other)(2012)[9]. In this paper, structure element descriptor (SED) – a novel texture descriptor, is proposed. SED can adequately describe images and represent image local features. Additionally, SED can extract and depict color and texture features. The image structure element's histogram (SEH) is processed by SED, and HSV color space is utilized (it has been quantized to 72 bins). SEH incorporates the advantage of both statistical and structural texture description techniques, and it can represent the spatial relationship of color and texture.

Greg Pass et al(and others)[10].In this paper color histograms are utilized to look at images in numerous applications. The advantages are efficiency, and insensitivity to small changes in camera perspective. They describe a histogram-based method for comparing images that incorporates spatial data. In this paper they have classify every pixel in a given color bucket as it is either coherent or incoherent, based on whether it is a part of a large similarity-colored region. A color coherence vector (CCV) contains the number of coherent pixels with every color. By isolating coherent pixels from incoherent pixels, CCV's give better distinctions than color histograms. They demonstrate that CCV's can give better results than color histograms for image retrieval.

Nikita Rishi et al(and other)(2015)[12].In this paper; heterogeneous plant diseases that are feasible and their

apprehension utilizing differentiating strategies have been examined. These strategies incorporate Otsu method, image compression, image cropping and image denoising including K means clustering to express the diseases images. Neural networks including back propagation (BP) networks, radial basis function (RBF) neural networks, a generalized regression networks (GRNNs) and probabilistic neural networks (PNNs) are used to analyse wheat and grape diseases. Cotton leaf diseases and rice plant diseases utilizing sobel operator, canny filter and feature extraction are passed down to recognize the diseases. Numerous different diseases like orchid leaf diseases, rubber tree leaf diseases; apple fruit diseases and chili plant diseases can likewise be experienced utilizing different methodologies like fuzzy logic, Multi-class Support Vector Machine(MSVM) and Local Binary Pattern(LBP). A smaller than usual explanation on every one of the diseases and their detection has been described in this paper.

O. Kleynen et al(and other)(2004)[11].In this paper a strategy to sort Jonagold apples relay on the presence of defects was proposed. A multi-spectral vision system incorporating four wavelength bands in the visible/NIR range was created. Multi-spectral images of sound and defected fruits were obtained tending to cover the entire color variability of this bicolor apple variety. Defects were gathered into four classifications: slight defects, more serious defects, defects leading to the rejection of the fruit and recent bruises. Stem-end/calyxes were identified using a correlation pattern matching algorithm. The accuracy of this strategy relied on upon the stem's orientation of the stem-end/calyx according to the optical axis of the camera. Defect categorization comprised in a pixel classification process based on the Bayes_ theorem and non-parametric models of the sound and defective issues. Fruit categorization tests were performed in order to measure the efficiency of the proposed system. No blunder was made on rejected products and high categorization rates were reached for apples presenting serious defects and recent bruises. Fruits with slight defects introduced a more imperative misclassification rate yet those blunders fitted however the quality tolerance of the European standard. Considering a real proportion of sound fruits of 90%, under 2% of infected fruits were classified into the sound ones.

Wen-Hung Liao et al(2010)[13].In this paper the local binary pattern (LBP) administrator is a computationally effective local texture descriptor and has discovered numerous helpful applications. Be that as it may, its sensitivity to noise and the high dimensionality of histogram connected with an ordinary size neighbourhood have raised a few concerns. In this paper, they attempted to upgrade the original LBP by proposing a novel expansion named extended local ternary pattern (ELTP). They will examine the qualities of ELTP in terms of noise sensitivity, discriminability and computational effectiveness. Preparatory trial results have demonstrated better adequacy of ELTP over the original LBP.

Basvaraj .S. Anami et al (and other)(2011)[14].In this paper an ANN based classifier is received which utilizes the combination of color and texture features to perceive and categorized different agriculture/horticulture produce. These features have given diverse accuracies in isolation for varieties of produce. The outcomes are empowering and guarantee the advancement of a good machine vision framework in the area of perceive and classification of farming/agriculture produce. The proposed methodology can fundamentally support in recognizing normal and affected produce.

The framework of the basic approach of the apple fruit diseases detection and classification is shown in the Fig. 3.

2.1. Image Segmentation Techniques
Segmentation means identifying region of interest from the image. In other word similar pixels are connected with each other [16].

The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyses [4]. Image segmentation methods are generally based on one of two fundamental properties of the intensity values of image pixels that is similarity and discontinuity.

2.2. Feature Extraction
Features are extracted from the fruit. Based on the features of the fruit the classification of fruit diseases is done. Extract various features from fruit like color of fruit, texture and shape[16].
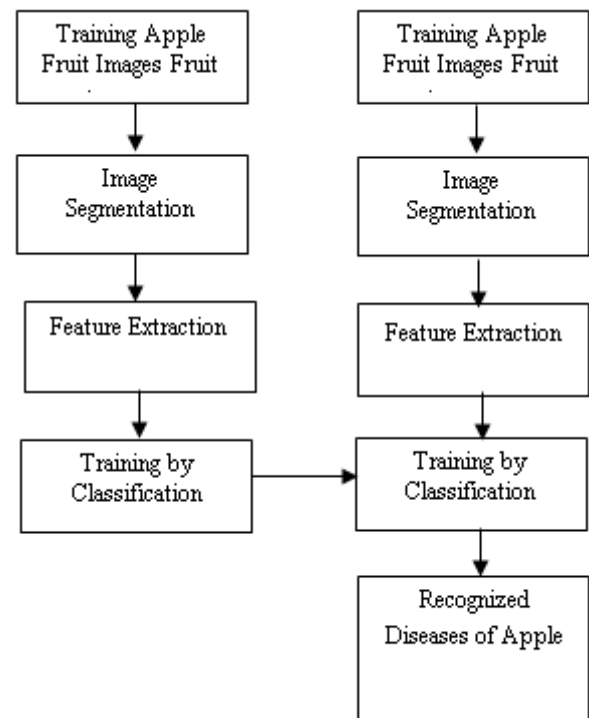

Fig 3: The Basic Procedure of the Apple Fruit Diseases Detection and Classification.

For detection of the fruit diseases Color and texture features play a vital role. Some color and texture features are extracted and used to accept the accuracy and efficiency of the framework. The color features utilized as a part of the fruit and vegetable classification/fruit diseases identification are HSV Histogram, L*a*b, RGB, YUV, Global Color Histogram, Color Coherence Vector, mean of three color array, Dominant color method, color difference histogram.

Table 1: Summary of different segmentation techniques[15]

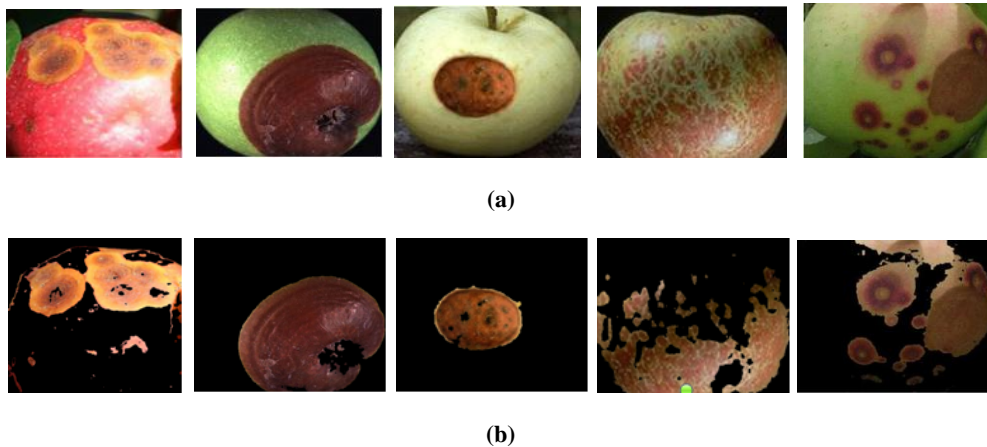| Segmentation Method	Description	Merits	Demerits
Histogram Matching	Histogram is constructed having peaks which relate to a region.	a) Low computation complexity. b) No prior information needed.	a) Spatial details not considered, cannot promise the segmented regions to be adjacent.
Region based approaches	Pixels are grouped in the homogeneous regions, and region merging, splitting or their combination is utilized.	a) Noise immune in edge detection approach.	a) High computational complexity. In region splitting segments seem square due to splitting scheme.
Edge detection approaches	Tries to locate the focuses having changes in gray level.	a) Works well for high contrast images.	a) Less immune to noise and doesn't work well if the image have many edges.
Clustering based approaches	Classify or cluster an image into several parts (regions) according to the feature of image.	a) Very simple. Reasonably fast.	a) Its "worst case" behavior is poor. b) The clusters are expected to be of similar size, so that the assignment to the nearest cluster center is the correct assignment.

(a)

(b)

Fig 4 : Some defect segmentation results on diseased fruit[1]

Table 2: Summary of different color feature techniques.

Color feature extraction method	Description	Merits	Demerits
HSV Histogram [11]	a) Hue, saturation value. b) It describes colors in terms of their shades and brightness. c) It is one of the main property (called color appearance parameter) of color.	a) High Accuracy. b) Suitable for real time application.	a) Less sensitive to lighting variations.
	a) Color information in L*a*b color space is stored only in	a) Color and intensity manage	a) Singularity problem as other

L*a*b [16]	two channels (i.e. a* and b* components). b) Space with dimension L for lightness and a and b for the color-opponent dimensions. c) It causes reduced processing time for the image segmentation.	independently. b) Ability to measure small color difference.	nonlinear transformation.
RGB[16]	a) An RGB color space is any additive color space based on the RGB color model. b)A particular RGB color space is defined by the three chromaticities of the red, green, and blue additive primaries. c)It can produce any chromaticity that is the triangle defined by those primary colors.	a) Convenient for display.	b) Due to its high correlation not good for color image processing.
YUV[16]	a) Y represents the luma component (the brightness) and U and V are the chrominance (color) components. b) YUV was invented when engineers wanted color television in a black-and-white infrastructure. c) They required a signal transmission strategy that was good with black and white (B&W) TV while being able to add color.	a) Less computation time and gets rid of correlation of RGB to some extent.	a) Correlation still exists but not as high as RGB.
Global Color Histogram[1]	a) Most straightforward deal with encode the information present in the image. b) It is an arrangement of ordered values, for each particular color, representing the probability of a pixel being of that color.	a) Simplest approach for distinct color.	a) Not more suitable.
Color coherence vector[1]	a) A way to deal with images based on color coherence vectors they characterize color coherence as the degree to which image pixels of that color are individuals of a large region with homogeneous color. These regions are called as coherent regions. b) Coherent pixels are fits to some sizable contiguous region, whereas incoherent pixels are not.	a) It is efficient and insensitive to small changes in camera viewpoint.	a) Color histograms lack spatial information, so image with very different appearance can have similar histograms.
Mean Of Three Color Array[16]	a) Color is calculated based on mean of three color value for red, green and blue.	a) Very easy for implementation.	a) Not accurate as other methods.

Table 3: Summary of different texture feature techniques.

Texture feature extraction method	Description	Merits	Demerits
Local Binary Pattern[1]	a) In which centre pixel and its corresponding neighbor pixels, calculate thresholding value for neighbor based on centre pixel.	a) It's robustness to monotonic gray-scale changes caused such as illumination variations. b)Its computational simplicity.	a) Binary data is sensitive to noise.
Complete Local Binary Pattern [1]	a) LBP features consider only signs of local differences. Whereas CLBP considers both signs (S) and magnitude (D) of local differences as well as original centre level (C) value.	a) It achieves much better rotation invariant texture classification results than conventional LBP-based schemes.	a) CLBP_S preserves more information of the local structure than CLBP_M, which shows why simple LBP operator can extracts texture features well.
Structure Element Histogram [8]	a) It is proposed to encode the small local structure of the small local structure of the images with its local features. b) It uses HSV color space. c) It is a combination of the advantages of both statistical and structural texture descriptor.	a) It provides a good symbolic description of the image. b) It is more useful for synthesis than analysis tasks.	a) The pixel number of the image will change when the image is scaled. b) Image would be dissimilar when the image is scaled.
Gray Level Co-occurrence Matrices [16]	a) It is in tabular form. In which how different combinations of pixel gray levels occur in an image. b) In which second order calculate relation between groups of two pixels.	a) Smaller length of feature vector. b) Applied it with different color space for color co-occurrence matrix.	a) They require a lot of computation (many matrices to be computed). b) It's not invariant with rotation and scaling
Gabor filter [16]	a) It is a signal processing method used for defining a set of radial centre frequencies and orientations.	a) It's a multi-scale, multi resolution filter. b) It has selectively for orientation, spectral bandwidth and spatial extent.	a) Large bank of filters used in application so computational cost is very high.
Wavelets Transform[16]	a) It is performed on the frequency domain rather than the spatial domain.	a) Produces best features with higher accuracy.	a) It is more complex and slower.
Independent Component Analysis[16]	a) It decomposes mixed signal into a set of linearly independent signals.	a)It is capable of obtaining higher order statistics. b)It is used to separate a multivariate signal implemented in texture classification.	a)It is new and not much popular method.

2.3. Classifiers

Classifier is used for classifying images based on their features[16]. There are many classifiers are available. K-Nearest Neighbours (k-NN), Support Vector Machine (SVM), Artificial Neural Network (ANN) and Random Forest Tree Classifier.

Table 4: Summary of different classifier.

Classifier	Description	Merits	Demerits
K-nearest neighbor [16]	a) It is a statistical classifier. b) It calculate distance metric for samples and classify based on this distance. c) It assigns data to the most represented category within its closest k neighbors. d) It uses Euclidean distance for distance calculation.	a) Simple implementation. b) Classification don't have to be linearly separable.	a) Sensitiveness to noisy or irrelevant data. b) Testing procedure is time consuming because of calculation of distance to all known instances.
Support Vector Machine [16]	a) It is also used for decision making. It works on two stages: 1) Off-line Process 2) Online Process b) Multi-class support vector machine as a set of binary SVM for training and classification.	a) SVM is well suited to work with high dimensional data. b) Classification accuracy is more as compared to other conventional classification techniques. c) SVM is robust enough, even when training samples have some distortion.	a) Section of kernel function and kernel parameters for mapping original data into higher dimensional data is difficult. b) Learning process can be time consuming. c) It works only with two classes appropriately.
Decision Tree [16]	a) It is a structure like a tree in which first of all source is split into subset based on its attribute value. b) Leaves represent class labels and branches represent conjunctions of features that prompt to those class.	a) Easy to interpret for small-sized trees. b) Accuracy is comparable to other classification techniques for many simple data sets.	a) This have been observed to over fit for some datasets with noisy classification/regression tasks.
Random Forest Algorithm [17]	a) It is an ensemble supervised machine learning technique. Based on bagging and random feature selection, number of decision trees (base classifiers) is generated and majority voting is taken for classification. b) The forest choose the classification having the most votes.	a) Deal really well with uneven data sets that have missing variables. b) It is one of the most accurate learning algorithms available. For many data sets, it produces a highly accurate classifier. c) It gives estimates of what variables are important in the classification. Faster than SVMs.	a) Random forests have been observed to over fit for some datasets with noisy classification/regression tasks.
Artificial Neural Network [16]	a) It is inspired by the concept of the biological neurons system. b) In which two dataset available one for training and another for testing.	a) Robust and user friendliness and can handle noisy data. b) Well suited to analyses complex number.	a) Scalability problem. b) Require large number training samples. c) Requires more processing time.

3. CONCLUSION

In this paper conclude that the different diseases detection methods for apple fruit diseases detection and distinctive color and texture techniques for feature extraction. Every techniques has some merits and negative marks. On the bases of requirement here need to utilize technique for color and texture. Additionally give outlines of distinctive classifier with its benefits and faults. Additionally see different segmentation techniques with its benefits and negative marks. Such segmentation method application in all images so utilize any of techniques which is suitable for our application. In future the apple fruit diseases detection and classification could be more accurate by fusing different color and texture features and the diseases classification using random forest algorithm.

4. REFERENCES

[1] Shiv Ram Dubey, Anand Singh Jalal," Detection and Classification of Apple Fruit Diseases using Complete Local Binary Patterns", 978-0-7695-4872-2/12 IEEE 2012.

[2] India Rank in Agriculture http://www.apeda.gov.in / apedawebsite/six_head_product/FFV.htm.

[3] P. Vimala Devi ,K.Vijayarekha," machine vision applications to locate fruits, detect defects and remove noise: a review" ,rasayan j.chem 2014.

[4] Shiv Ram Dubey, Pushkar Dixit, Nishant Singh, Jay Prakash Gupta,"Infected Fruit Part Detection using K-Means Clustering Segmentation Technique", International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 2, No 2,2013.

[5] R. Sivamoorthi, 2Dr. N. Sujatha," A Novel Approach of Detection and Classification of Apple Fruit Based on Complete Local Binary Patterns", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4, April 2015.

[6] Priya P.1, Dony A. D'souza," Study of Feature Extraction Techniques for the Detection of Diseases of Agricultural Products", international journal of innovative research in electrical, electronics, instrumentation and control engineering", Vol. 3, Special Issue 1, April 2015.

[7] Jagadeesh. D. Pujari, Rajesh. Yakkundimath ,A. S. Byadgi," Reduced Color and Texture features based Identification and Classification of Affected and Normal fruits' images", International Journal of Agricultural and Food Science 2013.

[8] Monika Jhuria,Ashwani kumar,Rushikesh Borse,"Image processing for smart farming: detection of diseases and fruit grading ", 978-1-4673-6101-9/13/2013 IEEE.

[9] Wang Xingyuan , Wang Zongyu," A novel method for image retrieval based on structure element's descriptor",Elsevier 2012.

[10] Greg Pass, Ramin Zabih,, Justin Miller," Comparing Images Using Color Coherence Vectors", Computer Science Department Cornell University.

[11] O. Kleynen , V. Leemans, M.-F. Destain," Development of a multi-spectral vision system for the detection of defects on apples",Elsevier 2004.

[12] Nikita Rishi, Jagbir Singh Gill," An Overview on Detection and Classification of Plant Diseases in Image Processing", International Journal of Scientific Engineering and Research (IJSER), Volume 3 Issue 5, May 2015.

[13] Wen-Hung Liao," Region Description Using Extended Local Ternary Patterns", 1051-4651/10 2010 IEEE.

[14] Basvaraj .S. Anami1, J.D. Pujari2, Rajesh.Yakkundimath," Identification and Classification of Normal and Affected Agriculture/horticulture ProducemBased on Combined Color and Texture Feature Extraction", International Journal of Computer Applications in Engineering Sciences, vol i, issue iii, september 2011.

[15] H.D. Cheng, X.H. Jiang, Y. Sun, Jingli Wang," Color image segmentation: advances and prospects", Elsevier 2001.

[16] Uravashi Solanki, Udesang K. Jaliya and Darshak G. Thakore ," A Survey on Detection of Disease and Fruit Grading", International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 2, 2015.

[17] Hossam M. Zawbaa,Maryam Hazman, Mona Abbass, Aboul Ella Hassanien,"Automatic fruit classification using random forest algorithm",978-1-4799-7633-1/14/2014 IEEE.