# Grading and Classification of Anthracnose Fungal Disease of Fruits based on Statistical Texture Features

Jagadeesh Devdas Pujari[1], Rajesh Yakkundimath[2]
and Abdulmunaf Syedhusain Byadgi[3]

*[1]S.D.M.College of Engg. &Tech, Dharwar – 580 008, India*
*[2]K.L.E.Institute of Technology, Hubli – 580 030, India*
*[3]University of Agricultural Sciences, Dharwar – 580005, India*
*[1]jaggudp@yahoo.com, [2]rajeshymath@gmail.com, [3]byadgi_uas@rediffmail.com*

***Abstract***

*In this paper, lesion areas affected by anthracnose are segmented using segmentation techniques, graded based on percentage of affected area and neural network classifier is used to classify normal and anthracnose affected on fruits. We have considered three types of fruit namely mango, grape and pomegranate for our work. The developed processing scheme consists of two phases. In the first phase, segmentation techniques namely thresholding, region growing, K-means clustering and watershed are employed for separating anthracnose affected lesion areas from normal area. Then these affected areas are graded by calculating the percentage of affected area. In the second phase texture features are extracted using Runlength Matrix. These features are then used for classification purpose using ANN classifier. We have conducted experimentation on a dataset of 600 fruits' image samples. The classification accuracies for normal and affected anthracnose fruit types are 84.65% and 76.6% respectively. The work finds application in developing a machine vision system in horticulture field.*

***Keywords:*** *Anthracnose fungal disease, Segmentation, RM texture features, Artificial neural network*

## 1. Introduction

India is the second largest producer of fruits with a production of 44.04 million tonnes from an area of 3.72 million hectares. This accounts 10% of the world fruit production. A large variety of fruits are grown in India of which apple, citrus, banana, grape, mango, guava, are the major ones. Also, India is a large low cost producer of fruit and horticulture has huge export potential. Especially in Karnataka state in India, that too around Dharwar, Bijapur, Bagalkot districts many farmers are depending on the fruits like mango, grape and pomegranate because of suitable climate condition and soil. Hence there is a major contribution from these areas in fruit production.

Fruit industry is a major industry which contributes 20% of the nation's growth. But due to improper cultivation of fruits, lack of maintenance and manual inspection there has been a decrease in production of good quality of fruits. Farmers are finding difficulty, especially in finding the fruits affected which results in huge loss of revenue to the farmers and the nation.

**Figure 1. Images of normal fruit produce**

**Figure 2.Images of anthracnose affected fruit produce**

Like any other crop, mango, grape and pomegranate is also susceptible to fungal disease, anthracnose. The emergence and spreading of diseases have become more common because of climate and environmental factors. The initial symptoms of anthracnose are characterized by small water soaked brownish spots in large numbers. Anthracnose disease attacks all plant parts at any growth stage like leaves, stems, and fruits. Symptoms are most visible on leaves, ripe fruits and stem. Image samples of both normal and anthracnose affected on fruit produce are shown in Figure 1 and Figure 2. Monitoring fruits plays a key role in successful cultivation. The naked eye observation is the main approach adopted so far.

These quality assessment procedures are carried out by experienced personnel. However, due to factors like different working conditions, personal judgment and level of fatigue, grading results often turn out to be inconsistent among individual inspectors. In future scenario, industry needs to secure its hard-won reputation in fruit quality and meanwhile increase fruit handling capacity of current facilities. To achieve this objective, the current visual inspection methods need to be automated. The fruit industry desires real-time quality evaluation tools capable of working consistently and objectively.

In the past few years, automation and intelligent sensing technologies have revolutionized our fruit production and processing routines. These initiatives have been accredited to the rising concerns about fruit quality and safety. Also, rising labor costs, shortage of skilled workers and the need to improve production processes have all put pressure on producers and processors. In such a scenario, automation can reduce the costs by promoting production efficiency. Automated solutions, such as quality grading and monitoring, post-harvest product sorting and robotics for field operations often integrate machine vision technology for sensing due to its non-destructive and accurate measurement capability.

To know the state-of-the-art in automation of the task/activities in horticulture field and automatic detection of fruit disease using computer vision techniques, a survey is made. The gist of a survey which carried out is given as follows.

(Dheeb Al Bashsish, *et al.*, 2011) have proposed a evaluate a software solution for automatic detection and classification of plant leaf diseases. The proposed detection based neural networks are very effective in recognizing leaf diseases, whilst k-means clustering technique provides efficient results in segmentation RGB images.

(Jayamala K. Patil1 and Raj Kumar, 2011) have provided advances in various methods used to study plant diseases/traits using image processing. The methods studied are for increasing throughput & reducing subjective ness arising from human experts in detecting the plant diseases. (Z. May and M. H. Amaran, 2011) have developed a new model of automated grading system for oil palm fruit is developed using the RGB color model and artificial fuzzy logic. The computer program is developed for the image processing part like the segmentation of colors, the calculation of the mean color intensity based on RGB color model and the decision making process using fuzzy logic to train the data and make the classification for the oil palm fruit. (Lili N.A., *et al.*, 2011) have used a modified Hierarchical Dynamic Artificial Neural Network which provides an adjustable sensitivity-specificity herbs diseases detection and classification from the analysis of noise-free colored herbs images. In this study, image processing and pattern classification are going to be used to implement a machine vision system that could identify and classify the visual symptoms of herb plants diseases. (D. S. Guru, *et al.*, 2011) have presented a novel algorithm for extracting lesion area and application of neural network to classify seedling diseases. First order statistical texture features are extracted from lesion area to detect and diagnose the disease type. These texture features are then used for classification purpose. A Probabilistic Neural Network (PNN) is employed to classify anthracnose and frog-eye spots present on tobacco seedling leaves. (D. Moshou, *et al.*, 2011) have developed a ground-based real-time remote sensing system that can be carried by tractors or robotic platforms is described. This prototype system makes possible the detection of plant diseases in arable crops automatically at an early stage of disease development and during field operations. The methodology uses differences in reflectance between healthy and diseased plants. Hyper spectral reflectance and multi-spectral imaging techniques were developed for simultaneous acquisition in the same canopy. (H. Al-Hiary, *et al.*, 2011) have evaluated a software solution for automatic detection and classification of plant leaf diseases. The proposed solution provides faster and more accurate solution removed. The experimental results demonstrate that the proposed technique is a robust technique for the detection of plant leaves diseases. (Anami B.S., *et al.*, 2009) have presented the use of computer vision technique on recognition and classification of bulk food grain image samples in the Indian context. C.C. Tuker S Chakraborty (2008) has presented software which detects and characterizes disease lesions on leaves to provide data on the number and type of lesions and the percentage of leaf area diseased using digital image processing (severity).

Most of the published work has mainly focused on disease detection on different crops. No attempt has been made on percentage of affected area and classification of anthracnose fungal disease from normal on fruits like mango, grape and pomegranate. Hence, it is the motivation for the present work to focus on image samples affected by anthracnose on fruits.

The technology leverage farmers can take up to asses the fruit, look at the possibility of diseases at early stages take decision on possible treatment and the like. In this paper, we have considered segmentation techniques to separate anthracnose affected lesion areas from normal area and graded based on percentage of affected area. A methodology is developed for determining whether, it is normal or affected by anthracnose. The chosen normal types are mango (Mauls domestic), grape (Vitas viniferous), pomegranate (Punic granite) and anthracnose affected types are amango, agrape, apomogranate. We have considered image samples normal and affected by anthracnose on ripe fruit, stem and leaf.

The paper is organized into four sections. Section 2 gives the proposed methodology. The results and discussions are given in Section 3. Section 4 gives conclusion of the work.

## 2. Proposed Methodology

In the present work, tasks like image acquisition, segmentation, feature extraction and classification are carried out. The detailed block diagram of adopted methodology is shown in Figure 3.

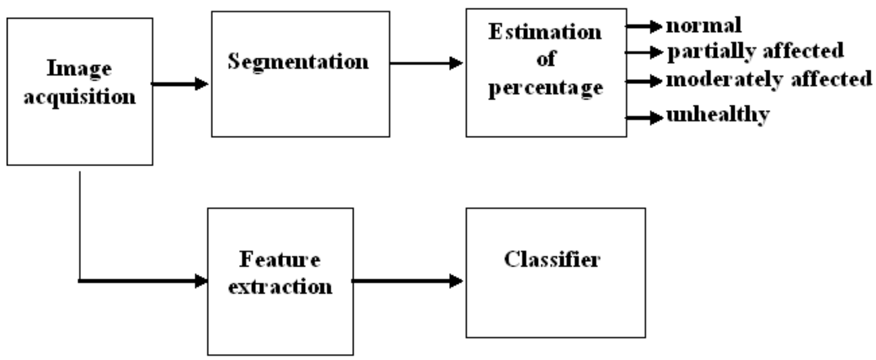



**Figure 3.Block diagram of proposed methodology**

### 2.1. Image Acquisition

Images of fruits' type both normal and affected by anthracnose are acquired with a color Digital Camera having a resolution of 12 mega pixels connected to a personal computer, Pentium IV, 2. 5 GHz. The images are taken keeping a distance of 0.5m from the samples.

### 2.2. Segmentation Techniques

In computer vision, segmentation refers to the process of clustering the pixels with certain properties into salient regions and these regions correspond to individual surfaces, objects or natural parts of the objects. More precisely, image segmentation is the process of assigning a label to every pixel in an image so that pixels with the same label share certain visual characteristics. The ultimate goal of segmentation is to simplify and/or change the representation of an image into a form more meaningful and easier to analyze. We propose the following techniques to segment the affected area.

**2.2.1. Thresholding:** Thresholding based segmentation is particularly effective for images containing objects resting upon a contrast background. During thresholding process, individual pixels in images are separates into background (binary "0") and foreground (binary "1") object-of-interest classes based upon their similarity in Gray-level intensity. The calculation of the actual threshold can be determined through ad-hoc experimentation. The grayscale image is converted into binary image using threshold (T1). A suitable threshold (T2) is used to remove small normal area i.e., the areas which are less than T2 pixels are removed. Figure 4(a) and (b), gives a sample input image and corresponding output image obtained by segmentation using thresholding technique.


(a)


(b)

**Figure 4 (a). Affected anthracnose leaf and (b) output of thresholding**

However, selecting a suitable threshold is a challenging task. If small threshold values are selected there are chances of retaining normal areas of small size and if large threshold values are selected there are chances of eliminating the affected areas. Hence, fixing up a suitable threshold T2 such that less probability for affected areas to be missing. Thresholding method is successful in most of the cases, provided the objects are to be clearly distinguishable from background and with the other classes of the object.

**2.2.2. Region growing:** In region growing approach, the neighboring pixels are examined to form a region or class, if no edges are detected. This process is iterated for each boundary pixel in the region. We have started with a seed pixel, examined local pixels around it and determined the most similar one based on the function of difference between neighboring pixels. The direction of region growth is chosen based on the neighboring pixel value differences. It is allowed to grow in the direction the difference in pixel intensity is less. This process is continued until no more pixels are added. Figure 5(a) and (b), gives a sample input image and corresponding output image obtained by segmentation using region growing technique.


(a)

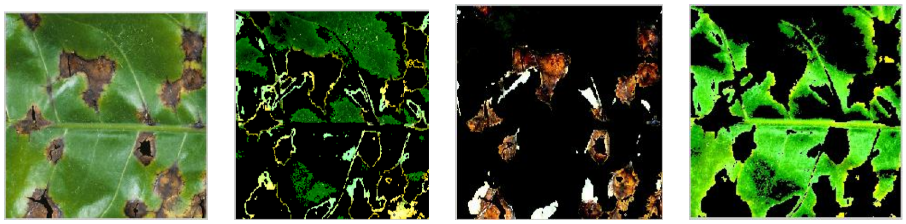

(b)

**Figure 5(a). Affected anthracnose leaf and (b) output of region growing**

The main drawback with region growing method it results in poor segmentation of fruit samples. The texture information is close to each other in similarly looking items like, healthy leaf images.

**2.2.3. K-means clustering:** The K-means clustering technique is a technique which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. In this method, k is the number of clusters in the segmented image. In our work, we have chosen k=3. The clustering is carried out based on the colors present in an image. We have used L*a*b* color space obtained from

RGB color space. Figure 6(a) gives a sample input image and Figure 6(b, c, d), gives corresponding first, second and third clusters respectively obtained by segmentation using K-means clustering technique.



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

**Figure 6(a). Affected anthracnose leaf and ((b), (c) & (d)) first, second and third clusters**

The main advantage with k-means clustering technique based segmentation is that it works on global information of the image in hand rather than local information. As the algorithm is computationally efficient, it is common to run it multiple times with different starting conditions to corroborate the results.

**2.2.4. Watershed:** A gray-level image may be seen as a topographic relief, where the gray level of a pixel is interpreted as its altitude in the relief. A drop of water falling on a topographic relief flows along a path to finally reach a local minimum. Intuitively, the watershed of a relief corresponds to the limits of the adjacent catchment basins of the drops of water.

For a segmentation purpose, the length of the gradient is interpreted as elevation information. In our work we have used marker controlled watershed segmentation technique. Local minima of the gradient of the image may be chosen as markers and involves region merging. Marker based watershed transformation make use of specific marker positions which have been either explicitly defined by the user or determined automatically with morphological operators or other ways. Figure 7(a) and (b) gives a sample input image and corresponding output image obtained by segmentation using marker controlled watershed segmentation technique.

(a) (b)

**Figure 7(a). Affected anthracnose leaf and (b) output of watershed**

The main drawback of marker controlled segmentation is computing internal and external markers which are used to modify gradient image. Computing markers depends on the specific nature of image samples.

**2.2.5. Estimation of percentage and grading:** To estimate the percentage of affected area on the fruits' image samples, we have separated the anthracnose affected lesion areas in an image using above segmentation techniques. Once lesion areas are separated from fruit image samples, the number of brown spots is counted to estimate the percentage of affected area.

The percentage of affected area is calculated using equation (1).

Percentage = (Affected area/Total area)*100 (1)

Where, Affected area= Number of brown spots count

Total area = Total size of leaf, stem or fruit

The affected area is calculated by counting number of brown spots over total area on leaf, stem and ripe fruits.

The grading is performed based on percentage of affected area, whether image sample is normal, partially affected, moderately affected or unhealthy as shown below.

Less than 1% affected area = Normal

Less than or equal to 25% affected area = Partially affected

Less than or equal to 50% affected area = Moderately affected

More than 50% affected area = Unhealthy

The procedure for estimating percentage of affected area is given in Algorithm 1. Table 1 gives number of image samples which are graded based on percentage affected across each segmentation techniques.

**Algorithm 1: Estimation of percentage of affected area**.

**Input:** Image samples
**Output:** Percentage of affected area
**Start**
Step1: Accept image samples both normal and affected by anthracnose
Step 2: Identify the area affected using segmentation techniques
Step 3: Estimate the percentage of affected area using Equation (1).
Step 4: If (percentage < 1)
Display ‘normal’
Else if (percentage<=25)
Display ‘partially affected
Else if (percentage<=50)
Display ‘Moderately affected’
Else
Display ‘Unhealthy’

**Stop.**

**Table 1.Estimation of grading based on percentage across segmentation techniques**

| Segmentation techniques | Grading | | | | |
|---|---|---|---|---|---|
| | Normal | Partially affected | Moderately affected | Unhealthy | Incorrect Estimation |
| Thresholding | 153 | 104 | 32 | 44 | 267 |
| Region growing | 97 | 66 | 20 | 49 | 368 |
| k-means clustering | 206 | 121 | 43 | 54 | 176 |
| Watershed | 107 | 74 | 29 | 33 | 357 |

### 2.3. Feature Extraction

The different fruit produce in spite of specific color, the color is not uniform over entire sample and changes over the size of the sample. First order statistics can be used as the most basic texture feature extraction methods, which are based on the probability of pixel intensity values occurring in digital images. Because of this fact, we have used the basic idea of Runlength statistics for extracting such information from gray level runs of an image. Consecutive pixels of the same gray value or level, in a given direction, constitute a run. The number of runs of different lengths and gray values form a two dimensional matrix called Runlength matrix. An element of RM, Q(x, y) represents the number of x gray values and y is considered run length. The procedure involved in generation of Runlength matrix in generation of Runlength matrix is given in Algorithm 2.

**Algorithm 2: Development of Run length Matrix $Q\varphi$ (x, y) from the Image *f(x, y)*.**

**Input:** Gray level image f(x, y) of size M*N

**Output:** Run length matrix $Q\varphi$(x, y) in the direction $\varphi$.

**Start**

Step 1: Assign $Q\varphi$(x, y) =0 for all x, y $\varepsilon$ [0, $L$], L is the maximum gray level.

Step 2: Find the matrix $Q\varphi$(x, y), for a given angle $\varphi$. The entry $Q$(x, y) is the (x, y)$^{th}$ entry in the Runlength matrix, where ‘x’ is the gray level and ‘y’ is the Runlength.

**Stop.**

We have obtained from the Runlength matrix, eight different features like entropy, short run, long run, Runlength Non-uniformity (RLN), Gray Level Non-uniformity (GLN), High gray Level Runlength Emphasis (HGRE), Low Gray Level Runlength Emphasis (LGRE) and Run Percentage (RP). These are used as features in the process of recognition and classification. We have found through experimentation that the features like entropy, short run and long run do not influence the recognition process. Hence, we have reduced the number of features to only five namely, runlength nonuniformity, gray level non-uniformity, low gray level runlength emphasis, high gray level runlength emphasis and run percentage.

These features are obtained using equations (2), (3), (4), (5) and (6) .The procedure adopted in texture feature extraction using Runlength matrix is given in Algorithm 3. The list of extracted texture features are shown in Table 2.

$$\text{RLN} = \sum_{j=1}^{R} \left( \sum_{i=1}^{G} Q\,(\mathrm{I}, j \mid \varphi) \right)^2 \qquad (2)$$

$$\text{GLN} = \sum_{i=1}^{G} \left( \sum_{j=1}^{R} Q\,(\mathrm{I}, j \mid \varphi) \right)^2 \qquad (3)$$

$$\text{LGRE} = \sum_{i=1}^{G} \sum_{j=1}^{R} Q\,(\mathrm{i}, \mathrm{j} \mid \varphi / (\mathrm{i}^*\mathrm{i}) \qquad (4)$$

$$\text{HGRE} = \sum_{i=1}^{G} \sum_{j=1}^{R} \mathrm{i}^2 Q\,(\mathrm{i}, \mathrm{j} \mid \varphi\,) \qquad (5)$$

$$\text{RP} = 1/\text{N} \sum_{i=1}^{G} \sum_{j=1}^{R} Q\,(i, j \mid \varphi\,) \qquad (6)$$

**Algorithm 3: RM Texture Feature Extraction**

**Input:** RGB components of original image
**Output:** Texture features
**Start**
Step 1: For all the separated RGB components perform Steps 2 thru Step 4.
Step 2: Derive the Run length Matrices $Q\varphi$(x, y) for four different directions φ ($0^0$, $45^0$, $90^0$ and $135^0$).
Step 3: Compute the Run length matrix, independent of direction using the equations (12) thru (16)
Step 4: Five Run length matrix features namely RLN, GLN, LGRE, HGRE and RP are calculated using equations (12) thru (16).
Stop.

**Table 2.RM Texture features**

| SL NO. | Features | SL NO. | Features | SL NO. | Features |
|---|---|---|---|---|---|
| 1 | Red RM RLN | 6 | Green RM RLN | 11 | Blue RM RLN |
| 2 | Red RM GLN | 7 | Green RM GLN | 12 | Blue RM GLN |
| 3 | Red RM LGRE | 8 | Green RM LGRE | 13 | Blue RM LGRE |
| 4 | Red RM HGRE | 9 | Green RM HGRE | 14 | Blue RM HGRE |
| 5 | Red RM RP | 10 | Green RM RP | 15 | Blue RM RP |

From the list of features given in Table 2 we have found that only 6 features contribute as discriminating features as this is essential for better classification. Hence we have considered only 6 features as second level feature reduction. The texture features reduced to 6 are listed in Table 3. Figure 8 shows reduced 6 texture features for each normal and affected fruit produce.

**Table 3.Reduced six RM Texture features**

| SL NO. | Features | SL NO. | Features | SL NO. | Features |
|---|---|---|---|---|---|
| 1 | Red RM RLN | 3 | Green RM RLN | 5 | Blue RM RLN |
| 2 | Red RM GLN | 4 | Green RM GLN | 6 | Blue RM GLN |

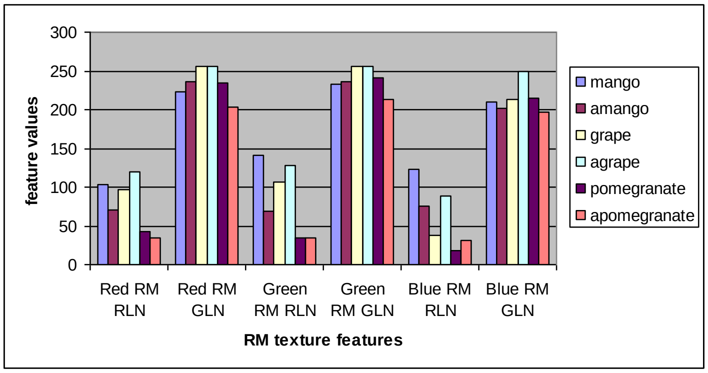



**Figure 8.Feature values based on six texture features for each fruit type**

### 2.4. Classifier

We have used a multilayered back propagation neural network (BPNN) as a classifier. The number of neurons in the input layer corresponds to the number of input features and the number of neurons in the output layer corresponds to the number of classes. The classifier is trained, validated and tested using images of different normal and anthracnose affected fruit type. We have kept the hidden layers to two arbitrarily. The developed neural network model performance is verified in terms of accuracy rate. The iterative reduction neural network model is analyzed. This indicates that as we reduce the number of redundant features from input layer accuracy reaches the maximum rate.

## 3. Results and Discussions

The MATLAB 7.0 tool box is used to implement the developed algorithms. We have considered 100 image samples of each normal type and 100 image samples of each anthracnose affected type amounting to a total of 600 image samples. Out of 300 anthracnose affected image samples 150 are of partially affected, 75 are of moderately affected and remaining 75 are of unhealthy.

For classification of image samples using ANN classifier, the network is trained with 80% images of each type. The remaining 20% images are used for testing. Around 15% of the image samples from the training set are used for validation of the designed classifier model.

The percentage accuracy is defined as the ratio of correctly recognized image samples to the total number of test image samples. The Percentage accuracy is given by equation.

$$\text{Percentage Accuracy} = \frac{\text{Correctly Recognised Image Samples}}{\text{Total Number of Test Image}} * 100$$

### 3.1 Grading of Image Samples

From Table 1 we have found that k-means based segmentation technique estimates correct percentage and grading for 424 image samples and 176 incorrect estimation. Region growing segmentation technique estimates correct percentage and grading for 232 image samples and 368 incorrect estimation. Thresholding based segmentation technique estimates correct percentage and grading for 333 image samples and 267 incorrect estimation. Watershed based segmentation technique estimates correct percentage and grading for 243 image samples and 357 incorrect estimation.

### 3.2. Identification Efficiency based on Reduced RM Texture Features

The reduced 6 RM texture features are extracted using Algorithm 2 and 3. The number of input nodes is 6 and number of output nodes is 6, for classifier. The classification accuracies of image samples for normal and anthracnose affected fruits' image samples is shown in Figure 9.

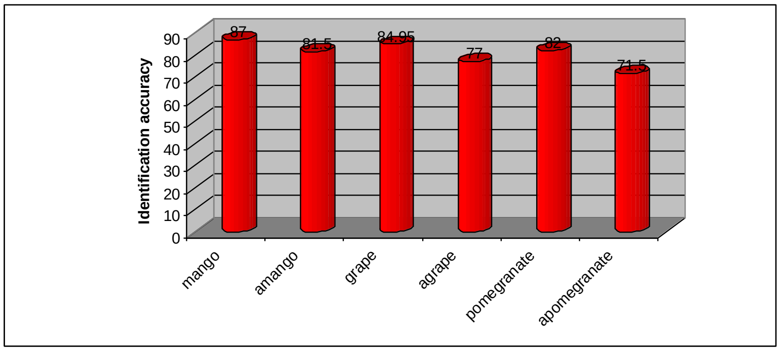



**Figure 9. Classification accuracy using reduced RM Texture features**

The highest recognition and classification accuracy of 87% is observed for normal mango and the lowest of 71.5% is observed with an anthracnose affected apomegranate.

### 3.3. Average Identification Efficiency based on Reduced RM Texture Features

The average accuracy of 84.65% for normal type and 76.6% for anthracnose affected type is achieved using reduced RM texture features. The average classification accuracies using reduced RM features are shown in Figure 10.

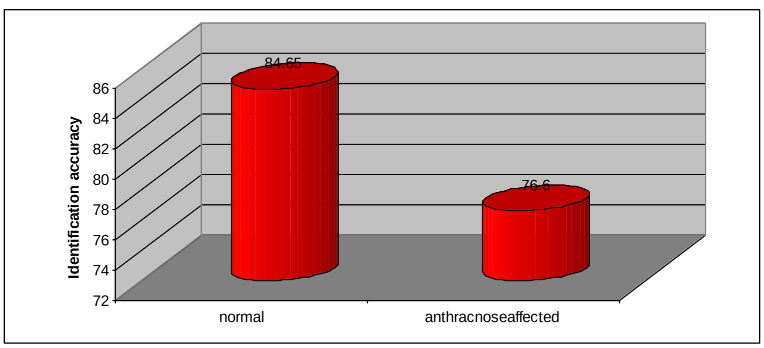



**Figure 10. Average classification accuracy using reduced RM Texture features**

## 4. Conclusions

The segmentation techniques like thresholding, region growing, K-means clustering, marker controlled watershed, are used to separate anthracnose affected lesion area from normal area. The percentage of affected area is estimated and graded among fruits' image samples. It is found that segmentation using K-means clustering founds to be better segmentation technique which estimates correct grading for 424 image samples and incorrect estimation for 176 image samples. Region growing supposed to be poor segmentation technique which estimates correct grading for 232 image samples and incorrect estimation for 368 image samples.

The RM texture features reduced from 15 to 6. The highest recognition and classification accuracy of 87% is observed for normal mango and the lowest of 71.5% is observed with an anthracnose affected apomegranate. The average accuracy of 84.65% for normal type and 76.6% for anthracnose affected type. A BPNN classifier is found suitable in this work.

The work carried out has relevance to real world classification of fruits' disease and it involves both image processing and pattern recognition techniques.

## References

[1] D. Al Bheesh, M. Braik and S. Bani-Ahmad, "Detection and Classification of leaf diseases using K-means-based segmentation and Neural-networks based Classification", Information Technology Journal, **(2011)**, pp. 265-267.

[2] J. K. Patil and R. Kumar, "Advances in image processing in image processing for detection of plant diseases", Journal of Advanced Bioinformatics Applications and Research, vol. 2, Issue 2, **(2011)**, pp. 135-141.

[3] Z. May and M. H. Amaran, "Automated Oil Palm Fruit Grading System using Artificial Intelligence", International Journal of Video & Image Processing and Network Security, vol. 11, no. 03, **(2011)**.

[4] N. A. Lili, F. Khalid and N. M. Borhan, "Classification of Herbs Plant Diseases via Hierarchical Dynamic Artificial Neural Network after Image Removal using Kernel Regression Framework", International Journal on Computer Science and Engineering, vol. 3, no. 1, **(2011)**.

[5] D. S. Guru, P. B. Mallikarjuna and S. Manjunath, "Segmentation and Classification of Tobacco Seedling Diseases", COMPUTE '11 Proceedings of the Fourth Annual ACM Bangalore Conference, **(2011)**.

[6] D. Moshou, C. Bravo, R. Oberti, J. S. West, H. Ramon, S. Vougioukas and D. Bochtis, "Intelligent multi-sensor system for the detection and treatment of fungal diseases in arable crops", Biosystems Engineering, vol. 108, **(2011)**, pp. 311-321.

[7] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and A. Rahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, vol. 17, no. 1, **(2011)**.

[8] B. S. Anami and D. G. Savakar, "Improved Method for Identification and Classification of Foreign bodies, mixed food grains", Image sample, ICGST/AIML Journal, vol. 9, **(2009)**, pp. 1-8.

[9] C. C. Tuker and K. Chakraborty, "Quantitative Assessment of Lesion Characteristics and Disease Severity Using Digital Image Processing", Journal of phypathology, vol. 145, Issue 7, **(2008)**, pp. 273 - 278.