Classification of Pomegranate Fruit using Texture Analysis of MR Images

A. Khoshroo[1*] , A.Keyhani[1], R.A.Zoroofi[2], S.Rafiee[1], Z. Zamani[3], M.R. Alsharif [4]
[1]Faculty of Biosystems Engineering, University of Tehran, Karaj, Iran
[2] Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[3]Department of Horticultural sciences, University of Tehran, Karaj, Iran
[4]Department of Information Engineering, University of the Ryukyus, Okinawa, Japan
[*]E-mail: Alireza_khoshroo@yahoo.com

ABSTRACT

Images obtained by Magnetic Resonance Imaging (MRI) of Iranian important export cultivar of pomegranate Malase-e-Torsh were analyzed by texture analysis to determine Gray Level Co-occurrence Matrix (GLCM) and Pixel Run-Length Matrix (PRLM) parameters. The T_2 slices measured at 1.5 T for 4 quality classes of pomegranate semi-ripe, ripe, over-ripe and internal defects classes were analyzed numerically using the software MaZda. To classify pomegranate into different classes, discriminant analysis was conducted using cross-validation method and texture features. Ten GLCM and 5 PRLM features were used in 2 different classifiers. Mean classification accuracy was 95.75 % and 91.28 % for GLCM and PRLM features respectively. By using GLCM and RPLM features, classification accuracy for semi-ripe, over-ripe and internal defects classes was higher when GLCM features were used. Ripe class had higher classification accuracy while PRLM features were used. To improve classification accuracy, combination of GLCM and PRLM features were used. For achieving best classification accuracy, optimum numbers of features were selected based on their contribution to the model. Combination of 7 GLCM and 4 PRLM features resulted in mean accuracy of 98.33 % and the lowest type I and II errors. Especially, type I error in ripe and over-ripe classes were significantly decreased. The classification accuracies were 100, 98.47, 100 and 95 % for semi-ripe, ripe, over-ripe and internal defects classes.

Keywords: Pomegranate, MRI, image processing, classification, texture, co-occurrence, run-length, discriminant analysis

1. INTRODUCTION

Pomegranate (*Punica granatum* L.) is an important fruit of tropical and subtropical regions. It is extensively cultivated in Iran, Mediterranean countries, India and to some extent in the U.S. (California), China, Japan and Russia. Iran is a native land of the pomegranate which is grown in every region, both coastal and mountainous areas (Fadavi *et al*., 2006). The pomegranate has been well known for its considerable pharmacological properties with anti-microbial, anti-viral, anti-cancer, potent anti-oxidant and anti-mutagenic effects (Negi *et al*, 2003). High-quality of product is the basis for success in today's highly competitive market. Harvest maturity influences quality and the nature of disorders during storage life of fresh fruits (Prabhu Desai, 1989). The pomegranate fruit has low respiration rate and non-climacteric respiratory pattern (Ben-Arie, 1984). Early harvest may impede the development of the characteristic colour, taste and aroma of pomegranates, while late-harvested fruits exhibit a reduced shelf life (Kulkarni *et al*., 2005).

At present, manual inspection is being used in order to determine quality of pomegranate. Increasing demand for quality assurance necessitate simple and reliable sorting methods. The use of computer vision system for quality assessment of pomegranate has also been reported (Khoshroo *et al*., 2006). Although these sorting methods succeed to a certain degree, they are limited by lack of ability to detect internal defects or internal quality. Product classification based on internal quality is almost non-existent in pomegranate packing houses and external quality is the main factor in the sorting lines. Hence products with good appearance and internal defects may pass undetected and may damage the surrounding healthy fruits during storage. A potential solution for detecting internal defects and maturity of pomegranate can be the use of non-destructive sensing technique. The ability of magnetic resonance imaging (MRI) to measure and quantify physical and chemical properties directly or indirectly provides a powerful tool for quality assessment (McCarthy, 1994). MRI has been shown to be an effective technique for internal quality assessment in a wide variety of fruit species including apple, avocado, kiwifruit, mango, melon, onion, orange, papaya, pear, peach, pineapple, potato, tangerine, tomato, strawberry, melon, and watermelon (Hills and Clark, 2003). Lammertyn *et al.* (2003) used MRI to monitor the development of core breakdown in 'Conference' pears during storage. They also reported that the contrast between affected and unaffected tissue was higher on the MR images in comparison with images from X-ray CT scans.

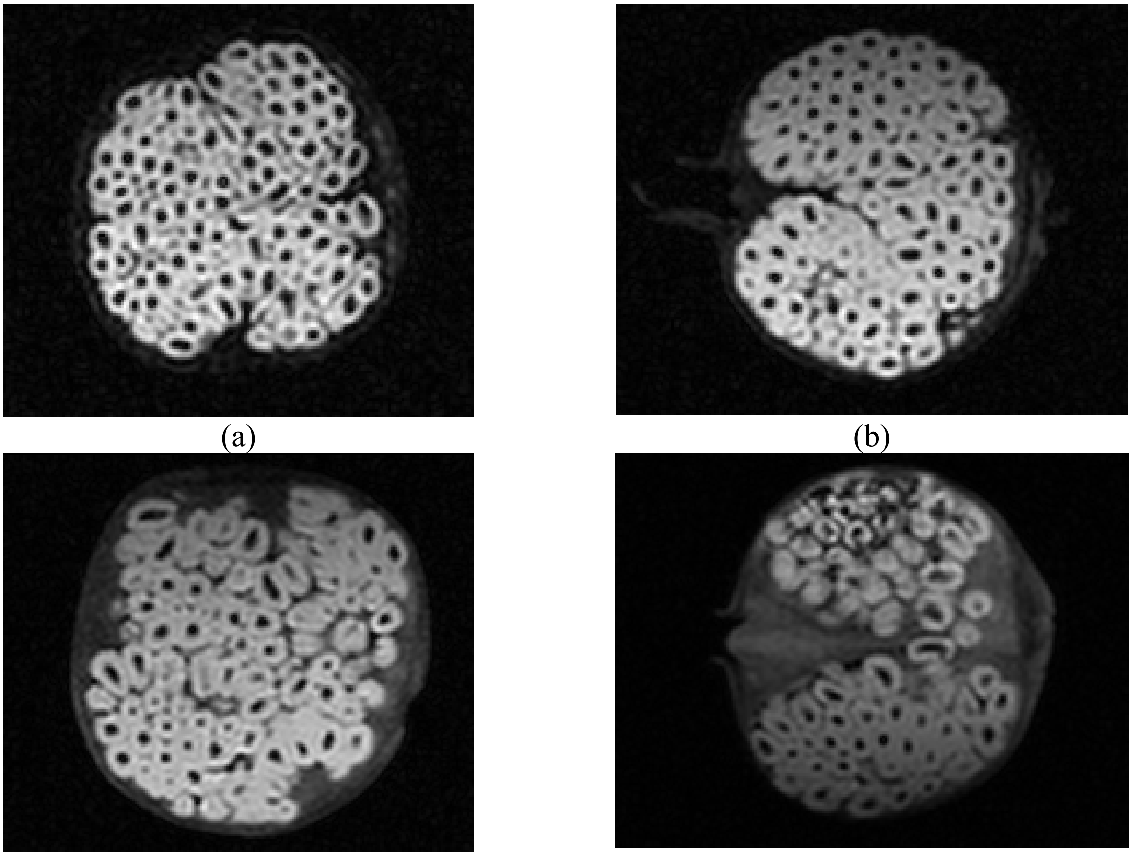
MRI can determine the changes in the internal texture of intact fruits during the fruit development and describes the internal distribution of affected tissues with high resolution (Clark *et al*., 1998). Thybo *et al*. (2004) could predict sensory texture attributes of cooked potatoes with nuclear magnetic resonance imaging. They suggest that MRI relates to the water distribution and some anatomic structures within the raw potatoes which are of importance for perceived textural properties of cooked potato. Letal *et al.* (2003) analyzed the magnetic resonance images of apples during ripening using textural features. Acidity was significantly correlated with sum average, sum variance and sum entropy.

The objective of this study was to investigate the applicability of MR imaging for assessment of changes in pomegranate maturity stages and detection of internal defects. For this purpose, development of a classifier based on textural features of MR images was evaluated.

2. MATERIALS AND METHODS

Pomegranates from the cultivar 'Malas-e-Torsh', an important export cultivar in Iran, were harvested at 3 maturity stages of semi-ripe, ripe and over-ripe from orchard of the pomegranate research station in Saveh (in the centre of Iran). Pomegranates with the internal defects were also chosen to evaluate internal quality of fruit. The selected fruits were picked off from different trees and stored under air temperature and transported for MRI measurement the next day after harvesting.

The experiments were performed on a 1.5T MRI scanner (Symphony, Siemens, Germany) with two-dimensional (2D) spin echo sequence, in the Noor Clinic, Tehran. The following parameters of T_2-weighted MR images were used: TR (Repetition Time) = 3910 ms, TE (Echo Time) = 60 ms, Field of View = 27.8 cm, slice thickness = 1.3 mm, interslice gap = 1.56 cm, number of slices = 20, matrix = 336×512. Figure 1 shows MR images of pomegranate at different quality stages.



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

Figure 1. MR images of pomegranate at different quality stages: (a) semi-ripe, (b) ripe, (c) over-ripe, (d) internal defects.

Texture analysis of region of interests in MR images was done using MaZda 2.11 software (Institute of Electronics, Technical University of Lodz, Poland). Ten Gray Level Co-occurrence Matrix (GLCM) derived parameters and 5 Pixel Run-Length Matrix (PRLM) based parameters were computed.

2.1. Gray Level Co-occurrence Matrix

One of the well-known statistical tools for extracting texture information from images is the gray level co-occurrence matrix. Originally introduced by Haralick et al. (1973), GLCM measures second-order texture characteristics which play an important role in human vision, and has been shown to achieve a similar level of classification performance. The GLCM of an Nx × Ny image, containing pixels with gray levels (0, 1,. . . , G-1) is a two-dimensional matrix P(k, l), where each element of the matrix represents the probability of joint occurrence of intensity levels *k* and *l* at a certain distance d and an angle θ. The co-occurrence matrix is normalized by dividing each entry of the matrix by a normalizing constant (C) that is the total number of pixel pairs in the image.

$$p(k,l) = \frac{P(k,l)}{C} \tag{1}$$

The following features were calculated from the normalized co-occurrence matrix for $d=1$ and four main directions (0°, 45°, 90° and 135°) and their mean value were calculated for further analysis.

Angular Second Moment

$$f_1 = \sum_{k} \sum_{l} p^2(k,l) \tag{2}$$

Contrast

$$f_2 = \sum_{j} j^2 (\sum_{k} \sum_{l} p(k,l)) \tag{3}$$

Sum of Squares

$$f_3 = \sigma^2 = \sum_{k} \sum_{l} (k - \mu^2) p(k,l) \tag{4}$$

$$\text{where } \mu = \sum_{k} \sum_{l} k.p(k,l)$$

Correlation

$$f_4 = \frac{\sum_{k} \sum_{l} (kl) p(k,l) - \mu^2}{\sigma^2} \tag{5}$$

Inverse difference moment

$$f_5 = \sum_{k} \sum_{l} \frac{p(k,l)}{1 + (k-l)^2} \tag{6}$$

Sum average

$$f_6 = \sum_{j=0}^{2G-2(G-1)} j.p_{x+y}(j) \tag{7}$$

$$\text{where } P_{x+y}(j) = \sum_{k} \sum_{l} p(k,l) \qquad k + l = j$$

Sum Variance

$$f_7 = \sum_{j=0}^{2G-2} (j - f_6)^2 p_{x+y}(j) \tag{8}$$

Sum Entropy

$$f_8 = -\sum_{j=0}^{2G-2} p_{x+y}(j) \log(p_{x+y}(j)) \tag{9}$$

Entropy

$$f_9 = -\sum_{k} \sum_{l} p(k,l) \log(p(k,l)) \tag{10}$$

Difference Variance

$$f_{10} = \sum_j p_{x-y}(j).(j - \sum_k k.p_{x-y}(k))^2 \quad (11)$$

where $p_{x-y}(j) = \sum_k \sum_l p(k,l) \qquad |k - l| = \mathrm{j}$

2.2 Pixel Run Length Matrix

The gray level run is a set of consecutive pixels having the same gray level value. The matrix elements q(j, k) specifies the number of times that the picture contains a run of length *k,* in a given direction, consisting of points having gray level *j.* The following features were extracted from four principal directions (0°, 45°, 90° and 135°) and their mean values were calculated for further analysis (Galloway, 1975).

Short run emphasis

$$SRE = \frac{\sum_k \sum_l \frac{R(k,l)}{l^2}}{\sum_k \sum_l R(k,l)} \quad (12)$$

Long run emphasis

$$LRE = \frac{\sum_k \sum_l l^2 R(k,l)}{\sum_k \sum_l R(k,l)} \quad (13)$$

Gray level non-uniformity

$$GLNU = \frac{\sum_k (\sum_l R(k,l))^2}{\sum_k \sum_l R(k,l)} \quad (14)$$

Run-length non-uniformity

$$RLNU = \frac{\sum_l (\sum_k R(k,l))^2}{\sum_k \sum_l R(k,l)} \quad (15)$$

Fraction (run percent)

$$RP = \frac{\sum_k \sum_l R(k,l)}{\sum_k \sum_l lR(k,l)} \quad (16)$$

2.3 Discriminant Analysis

For data discrimination, the criteria to evaluate the effectiveness of features should be a measure of the class separability. Discriminant analysis, based on a family of functions of scatter considers a within-class scatter matrix for each class, measuring the scatter of samples around the respective class mean, and the between-class scatter matrix, measuring the scatter of class means around the mixture mean, and finds a transformation that maximizes the between-class scatter and minimizes the within-class scatter, so that the class separability is maximized in the reduced dimensional space (Sun, 2007). In this study, discriminant analysis was conducted using cross-validation method with normal estimation using SAS software (SAS Institute Inc., USA).

3. RESULTS AND DISCUSSION

In order to evaluate changes in texture in different quality classes, textural features of pomegranate MR images were extracted. GLCM features, PRLM features and combination of GLCM and PRLM features were used for classification of pomegranates by using discriminant analysis. Type I and II errors are the most common statistics to evaluate the performance of a classification procedure in pattern recognition (Ott and Longnecker, 2001). Type I error is the error rate of missing classified samples in each class. For example, in the ripe class, the type I error is the number of missing classified ripe pomegranate divided by the total number of ripe fruit. Type II error is the error rate of false classified samples in each class. For example in the ripe class, the type II error is the number of false classified ripe pomegranate divided by the total number of ripe fruit.

3.1. Model 1 - GLCM Features

In this model, the GLCM features were used. Table 1 shows the textural features extracted form the co-occurrence matrix in the descending order of their level of contribution. The *correlation* was the most significant feature (ASCC=0.164) and the *difference variance* was the least significant (ASCC=0.64).

Table 1. Selection of GLCM features of pomegranate MR images based on their contribution to the texture model

No.	GLCM textural features	Average Squared Canonical Correlation (ASCC)	Partial r^2
1	Correlation	0.164	0.49
2	Sum Average	0.241	0.33
3	Sum of Squares	0.403	0.56
4	Sum Entropy	0.459	0.33
5	Angular Second Moment	0.508	0.22
6	Contrast	0.528	0.20
7	Entropy	0.548	0.17
8	Difference Entropy	0.613	0.40
9	Inverse Difference Moment	0.632	0.14
10	Difference Variance	0.640	0.10

Table 2 displays classification confusion matrix of pomegranate using GLCM textural features and discriminant analysis. Classification accuracy for semi-ripe, ripe, over-ripe and internal defects class was 100 %, 93.13 %, 94.89 % and 95 %, respectively. The mean accuracy that is the average classification of 4 classes was 95.75 %.

Table 2. Classification confusion matrix of pomegranate sing GLCM features: Normal estimation (Cross-validation method)

Class	Semi-ripe	Ripe	Over-ripe	Internal Defects
Semi-ripe	80 (100%)	0	0	0
Ripe	0	122 (93.13%)	6	3
Over-ripe	0	7	130 (94.89%)	0
Internal Defects	1	1	0	38 (95%)

Type I and II errors are shown in figure 2. Ripe class had the highest type I error but internal defects class had the highest type II error.

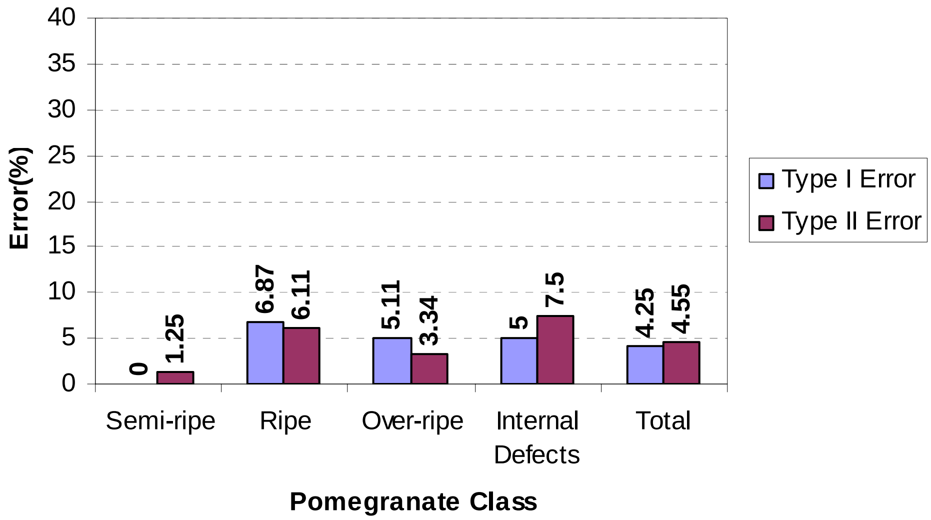


Figure 2. Type I and II errors for different pomegranate classes using GLCM features

3.2. Model 2 - PRLM Features

In the second model, PRLM features were used for classification of pomegranates into different quality classes. Table 3 shows the PRLM parameters based on their level of contribution to the classification model. *Run length non-uniformity* (ASCC=0.134) and *gray level non-uniformity* (ASCC=0.215) were the most significant features.

Table 3. Selection of PRLM features of pomegranate MR images based on their contribution to the texture model

No.	PRLM textural features	Average Squared Canonical Correlation (ASCC)	Partial r^2
1	Run length non-uniformity	0.134	0.40
2	Gray level non-uniformity	0.215	0.25
3	Short Run Emphasis	0.302	0.34
4	Fraction	0.357	0.26
5	Long Run Emphasis	0.408	0.19

Table 4 demonstrates classification confusion matrix of pomegranate using PRLM features. Classification accuracy for semi-ripe, ripe, over-ripe and internal defects class was 95 %, 94.66 %, 83.95 % and 87.5 % respectively. The mean accuracy was 91.28 %.

Table 4. Classification confusion matrix of pomegranate using PRLM features: Normal estimation (Cross-validation method)

Class	Semi-ripe	Ripe	Over-ripe	Internal Defects
Semi-ripe	76 (95%)	1	1	2
Ripe	2	124 (94.66%)	0	5
Over-ripe	13	12	115 (83.95%)	7
Internal Defects	2	3	0	35 (87.5%)

Type I and II errors are shown in figure 3. In general, PRLM features showed higher type I and II error than that of GLCM features.

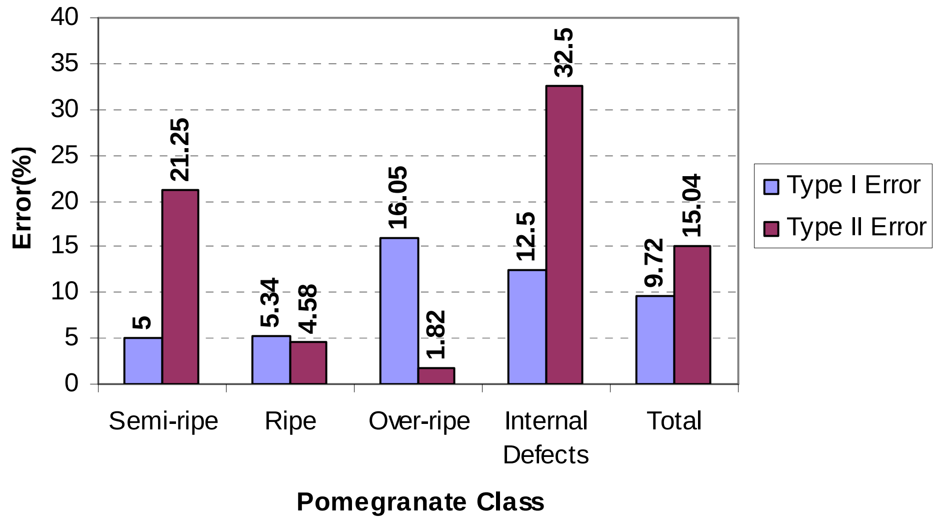


Figure 3. Type I and II errors for different pomegranate classes using PRLM features

3.3. Model 3 - GLCM+PRLM Features

In order to improve the classification of pomegranates, in the third model, combination of GLCM and PRLM features were used. The parameters with the highest discriminative power are shown in the descending order in table 5. *Correlation* and *Gray level non-uniformity* were the most significant features. Discriminant analysis was performed with different number of features and classification accuracies were compared (Figure 4). The mean accuracy was found to be poor when only the first five features were used. As the number of features increased, the mean accuracy increased to a certain extent and then decreased due to redundancy of some features in the model. The Classification model with 11 most significant features gave a satisfactory classification accuracy of 98.33 %.

Table 5. Selection of GLCM and PRLM textural features of pomegranate MR images based on their contribution to the texture model

No.	GLCM and PRLM textural features	Average Squared Canonical Correlation (ASCC)	Partial r^2
1	Correlation	0.164	0.49
2	Gray Level Non Uniformity	0.282	0.35
3	Sum Average	0.381	0.51
4	Difference Variance	0.473	0.43
5	Sum Entropy	0.532	0.37
6	Long Run Emphasis	0.555	0.25
7	Angular Second Moment	0.602	0.25
8	Short Run Emphasis	0.627	0.18
9	Entropy	0.643	0.14
10	Difference Entropy	0.682	0.25
11	Fraction	0.692	0.12
12	Run length non-uniformity	0.698	0.09
13	Sum Of Squares	0.706	0.07
14	Contrast	0.717	0.07
15	Inverse Difference Moment	0.725	0.06

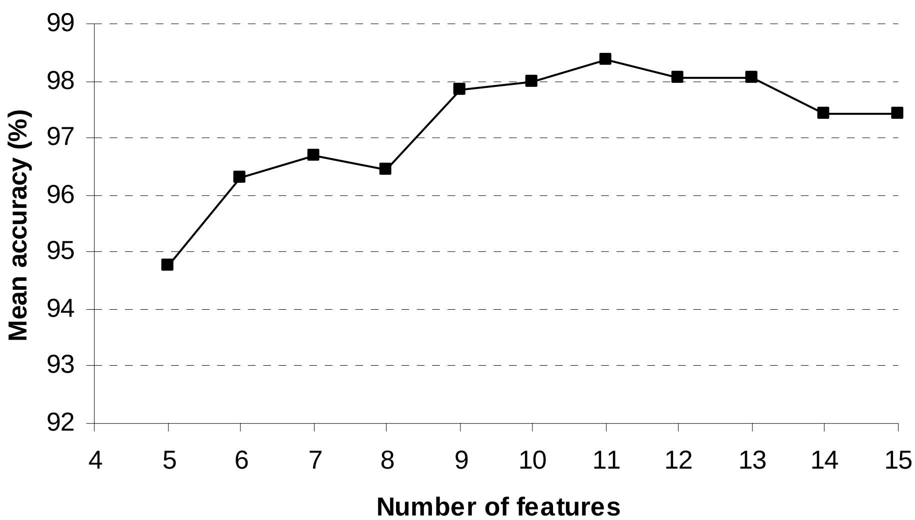


Figure 4. Classification accuracies of pomegranate classes using GLCM and PRLM textural features

Table 6 demonstrates classification confusion matrix of pomegranate using 11 most significant GLCM and PRLM features. Classification accuracy for semi-ripe, ripe, over-ripe and internal defects class was 100 %, 98.47 %, 100 % and 95 %, respectively.

Type I and II errors are shown in figure 5. Internal defects class showed the highest Type I error while ripe class had the highest type II error. In comparison with other models, this model had the lowest Type I and II errors.

Table 6. Classification confusion matrix of pomegranate using textural features: Normal estimation (Cross-validation method)

Class	Semi-ripe	Ripe	Over-ripe	Internal Defects
Semi-ripe	80 (100%)	0	0	0
Ripe	0	129 (98.47%)	1	1
Over-ripe	0	0	137 (100%)	0
Internal Defects	1	1	0	38 (95%)

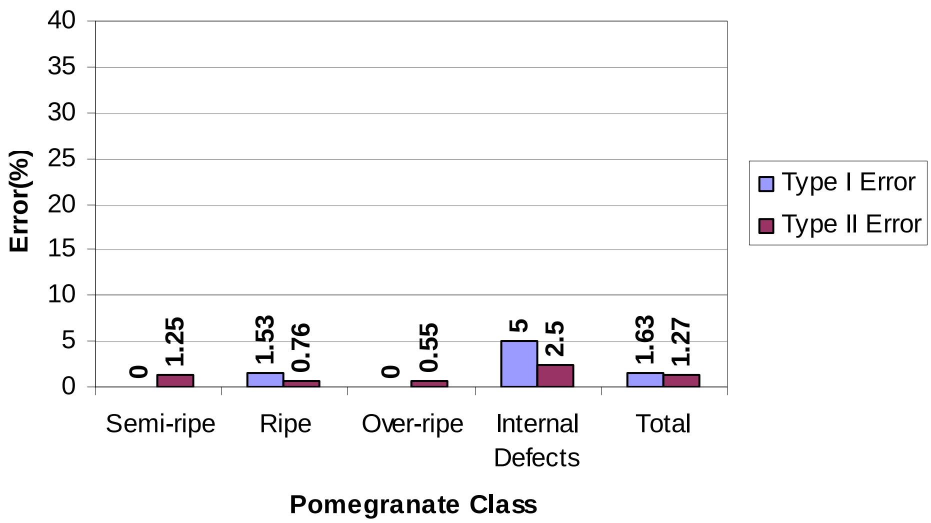


Figure 5. Type I and II error using GLCM and PRLM Features

Comparison of 3 classifiers in classification of pomegranate into different quality classes using MR images showed that combination of GLCM and PRLM features resulted in high classification accuracy of 98.33 %. This confirms that MRI is a powerful tool in visualizing internal structure of pomegranate fruit. The MRI technique is able to discriminate between healthy and defective fruits with high accuracy. Also, MRI can provide valuable information about internal changes in pomegranate during different maturity stages. Since maturation process may result in increase in free water and MRI is able to detect changes in amount and distribution of water, MRI can detect maturity in pomegranate fruits. Free water in over-ripe pineapple flesh and contrast differences in flesh of green and ripe tomato and avocado were also clearly detected by MRI (Chen et al., 1989).

4. CONCLUSIONS

Magnetic resonance imaging was used to visualize internal structure of pomegranate. To determine ripening stage and internal quality of fruit, texture analysis of MR images was performed. Using co-occurrence features for classification, the mean accuracy was 95.75 %. Type I error was found to obtain the highest value in the *ripe* class while type II error showed the highest value in *Internal Defects* class. Classification accuracy was decreased to 91.28 % when pixel run-length features were individually used.

Combination of co-occurrence and pixel run-length matrix features resulted in decreasing type I and II errors. Combination of 7 co-occurrence features and 4 pixel run-length features resulted in the mean accuracy of 98.43 % and the lowest type I and II errors. Especially type I error in ripe and over-ripe pomegranates was significantly decreased.

5. REFERENCES

Ben-Arie, R., N. Segel and S. Guelfat-Reich. 1984. The maturation and ripening of the wonderful pomegranate. *Journal of the American Society for Horticultural Science* 109(6): 898–902.

Clark, C. J., J. S. MacFall and R. L. Bieleski. 1998. Loss of watercore from ‘Fuji’ apple observed by magnetic resonance imaging. *Scientia Horticulturae* 73(4): 213–227.

Fadavi, A., M. Barzegar and M. H. Azizi. 2006. Determination of fatty acids and total lipid content in oilseed of 25 pomegranates varieties grown in Iran. *Journal of Food Composition and Analysis* 19(6): 676-680.

Galloway, M. M. 1975. Textural analysis using grey level run lengths. *Computer Graphic and Image Processing* 4: 172–179.

Haralick, R. M., K. Shanmugam and I. Dinstein. 1973.Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* 6: 610–621.

Hills, B. P. and C. J. Clark. 2003. Quality assessment of horticultural products by NMR. *Annual Reports on NMR Spectroscopy* 50: 75–120.

Khoshroo, A., R. A. Zoroofi, A. Keyhani, Z. Zamani, and S. Rafiee. 2006. Pomegranate quiality evaluation using machine vision. Abstract book of International Symposium on Pomegranate and Minor Mediterranean Fruits. Adana, Turkey.

Kulkarni, A. P. and S. M. Aradhya. 2005. Chemical changes and antioxidant activity in pomegranate arils during fruit development. *Food Chemistry* 93: 319–324.

Lammertyn, J., T. Dresselaers, P. Van Hecke, P. Jancsok, M. Wevers and B.M. Nicolai. 2003. MRI and X-ray CT study of spatial distribution of core breakdown in ‘Conference’ pears. *Magnetic Resonance Imaging* 21: 805–815.

Letal, J., D. Jirak, L. Suderlova and M. Hajek. 2003. MRI texture analysis of MR images of apples during ripening and storage**.** *Lebensmittel-Wissenschaft und -Technologie* 36: 719–727

MaZda 2.11 (1999). MRI Analysis Software, Technical University of Łodz, Poland, P. Szczypinski.

McCarthy, M. J., 1994. *Magnetic Resonance Imaging in Foods*. Chapman & Hall, New York.

Negi, P. S., G. K. Jayaprakasha, and B. S. Jena. 2003. Antioxidant and antimutagenic activities of pomegranate peel extracts. *Food Chemistry* 80: 393–397.

Ott, R. L. and M. Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis.* 5th ed. Duxbury Press, Pacific Grove CA.

Prabhu Desai, V.G. 1989. *Investigation on internal breakdown of pomegranate fruit*. Ph.D. Thesis, MPWV, Rahuri.

SAS. 1990. SAS User's Guide: Statistics. Raleigh, N.C.: Statistical Analysis System, Inc.

Sun, D. W. 2007. *Computer Vision Technology for Food Quality Evaluation.* Academic Press.

Thybo, A. K., S. N. Jespersen, P. E. Lærke and H. J. Stødkilde-Jørgensen. 2004. Nondestructive detection of internal bruise and spraing disease symptoms in potatoes using magnetic resonance imaging. *Magnetic Resonance Imaging* 22:1311–1317.