

## POMEGRANATE FRUIT QUALITY ASSESSMENT USING MACHINE INTELLIGENCE AND WAVELET FEATURES

Arun KUMAR R.\*, Vijay S. RAJPUROHIT, Bhairu J. JIRAGE  
 KLS Gogte Institute of Technology  
 Udyambag  
 590008 Belagavi, India

Received: March 2018; Accepted: June 2018

### ABSTRACT

Quality assessment is an important concern in the post-harvest marketing of fruits. Manual quality assessment of pomegranate fruits poses various problems because of human operators. In the present paper, an efficient machine vision system is designed and implemented in order to assess the quality of pomegranate fruits. The main objectives of the present study are (1) to adopt a best pre-processing module, (2) to select best class of features and (3) to develop an efficient machine learning technique for quality assessment of pomegranates. The sample images of pomegranate fruits are captured using a custom-made image acquisition system. Two sets of features, namely, spatial domain feature set and wavelet feature set are extracted for all of the sample images. Experiments are conducted by training both artificial neural networks (ANNs) and support vector machines (SVMs) using both sets of features. The results of the experiments illustrated that ANNs outperform SVMs with a difference in the accuracy of 12.65%. Further, the selection of wavelet feature set for training yielded more accurate results against spatial domain feature set.

**Key words:** Pomegranate quality assessment, wavelet features for fruit quality analysis, machine intelligence in pomegranate quality assessment, ANN for fruit quality assessment

### INTRODUCTION

Pomegranate (*Punica granatum*) has emerged as an important key fruit crop of semi-arid and arid regions. This is because it has immense medicinal value, gives high profit with low investments, has an ability to survive long-distance transport along with a built-in capacity to resist heat, drought and moisture deficit.

There is a huge potential for exporting pomegranates from India. The scope of export of Indian pomegranates is stretched to Qatar, Bangladesh, Bahrain, Saudi Arabia, Canada, Germany, United Kingdom, Japan, Kuwait, Sri Lanka, Oman, Pakistan, Singapore, Switzerland, U.A.E. and U.S.A. Exports from India have grown by 5% over the previous seven years (APEDA 2015).

Unfortunately India produces only 2% of the world's total amount of pomegranate. Moreover there are no disciplined marketing systems for testing

the quality of pomegranates (Babu 2012). Farmers marshal their fruit produce to the contractors who are then responsible for transporting it to distant markets (Benagi et al. 2009). Hence there is a vital need to propose a working system for quality assurance of pomegranate fruits post-harvest.

A computer vision system includes the application of techniques in which computers are employed to examine and extract image contents in solving specific problems concerning the fruit surface. Quality assessment of agricultural produce offers definite challenges as the “appearance” is inconsistent and vague (Deepa & Geethalakshmi 2011). Food industry is amid the top ten industries that widely use the machine vision. Its role is exceptional in the field of automated qualification and sorting of horticultural, agricultural and food products.

The present study describes development of a machine vision system to assess pomegranate fruits into two classes, viz. diseased and healthy.

\*Corresponding author:  
 e-mail: kumararun37@gmail.com

The methodology starts with applying histogram equalization of the captured fruits followed by wavelet denoising. The pre-processed images are then fed to a feature extraction algorithm that extracts spatial domain features and wavelet features of each of the image. The features are then fed to two classifiers, namely, artificial neural networks (ANN) and support vector machines (SVM). Experiments are conducted to determine (a) the best class of feature and (b) best classifier.

The major contributions of the current study are as follows: (1) Histogram equalization followed by the application of wavelet denoising. This method can in general be extended to any image preprocessing application. (2) Three of the wavelet families are applied to wavelet transforms of the images, each at three levels of subdivision.

The automated qualification and sorting of agricultural produce has been the important subject for various researchers across the globe. Teimouri et al. (2014) developed an efficient algorithm for segmenting almond images using artificial neural networks (ANNs) and obtained an overall accuracy of 98.82%. Nectarine variety was investigated and verified by Font et al. (2014) with an accuracy of 87%. Cucumbers were classified as per the European Grading Standards by Clement et al. (2013) with 99% accuracy. Narrow and broad weed were classified based on DWT features by Ghazali et al. (2007) with an accuracy of 87.25%. Palm oil fresh fruit bunches were automatically graded by Jamil et al. (2009) with the help of neuro fuzzy systems to an extent of 73.3%. Raisins were automatically graded by Omid et al. (2010) that achieved a classification rate of 96%. Multi-class fruits/vegetables were classified by Rocha et al. (2010) by combining features and classifiers, which resulted in an error reduction rate of 15%. A vision algorithm was developed by Arefi et al. (2011) to assist the robotic arm to pick mature tomato, which demonstrated an accuracy of 96.36%. Cucumber leaf diseases were classified with the help of computer image processing and support vector Machine (SVM) by Youwen et al. (2008) and obtained satisfying results when compared to that of ANNs. A new technique to sort and automatically grade fruits was devised by Mustafa

et al. (2009) with the help of fuzzy logic and obtained the promising results.

Because of the very few studies carried out on sorting of pomegranate fruits by machine vision, we found that there is a large scope for pomegranate sorters at the industry level. This is especially needed for use at the fruit auction centers, where sorting of pomegranate fruits is still done manually. The present work is aimed at developing a machine vision system to sort the pomegranate fruits into two quality categories.

## MATERIALS AND METHODS

The overall methodology of the present work is depicted in Fig. 1. Each subsection discusses each of the modules in detail. The entire development of the project is done using Matlab R2017a.

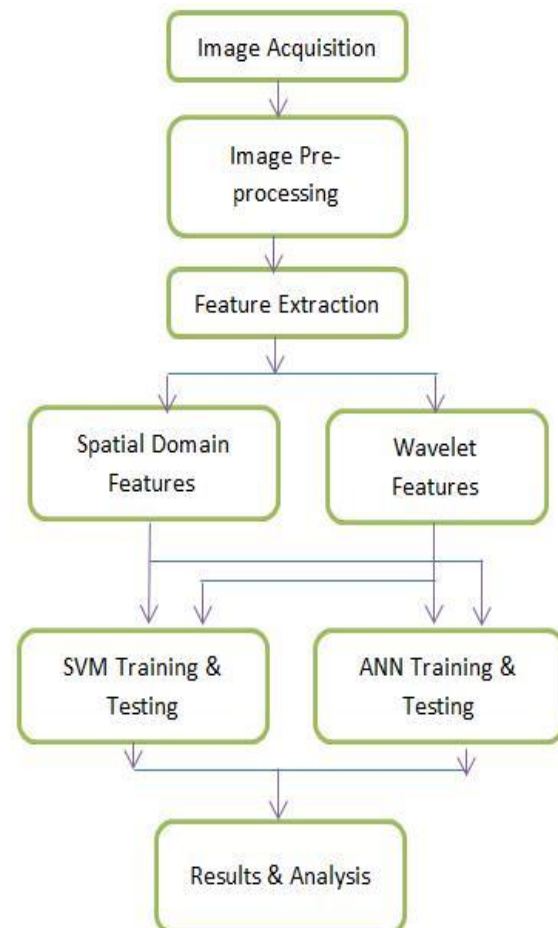


Fig. 1. Overall methodology of pomegranate fruit sorting

### Image acquisition

A closed metal compartment is built for the purpose of capturing images of pomegranate fruits. There is a provision for mounting light source and cameras that mimics the packing lines in fruit sorting industries. The image acquisition compartment is shown in Fig. 2. Table 1 summarizes the details of the acquisition system and settings, and plant material. A total of 60 diseased fruits and 60 healthy fruits were considered for the experiments. Fig. 3 depicts sample images of each class.

### Image preprocessing

Since images are captured using a light source, light carries an impact on the extraction of feature values out of the images. Therefore, we need to neutralize the effect of light. Technically, we need to equalize the spread of image intensity values. The equivalent

digital image operation is histogram equalization. Hence in the current work histogram equalization was initially applied to all the captured images. But one drawback of histogram equalization is that it reduces the informational content of the image. Therefore we need to make up for the information loss of the image. This can be achieved by applying wavelet denoising to the histogram-equalized images. Wavelet denoising endeavors to remove the noise present in the signal while conserving the characteristics of the signal, irrespective of the frequency content (Hazra & Guhathakurta 2016). This will radically increase the informational entropy of the image.

### Feature extraction

The preprocessed images are then separately fed to two modules: spatial features extraction module and wavelet features extraction module.

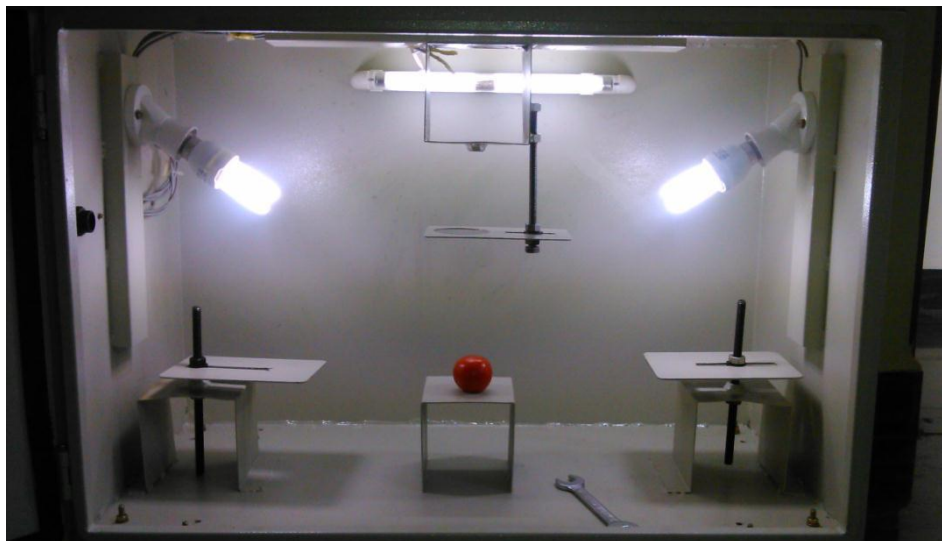


Fig. 2. Compartment for image acquisition

Table 1. Details of experimental materials

S1.No.	Particulars	Details
1	Light source	Compact Fluorescent Light Voltage: 240 V, Frequency: 50 Hz, Current: 65 mA, Power factor: 0.85, 765 Lumen
2	Camera	Logitech C905 720p Webcam, 2MP sensor Focal length: 3.7 mm, lens aperture f/2.0, focus adjustment: automatic, auto-exposure mode
3	Image resolution	1600 × 1200, 96 dpi (3.779528 pixel/mm)
4	Sample images	Cultivars: Bhagwa, Ganesh Source: Local Fruit Market (Belagavi, Karnataka, India) Number of days after harvest: 6 Selection of healthy and diseased samples: Based on human expert opinion



Fig. 3. Sample diseased class and healthy class

Table 2. Spatial domain features

Sl. No.	Feature Name	Remark
Image region features		
1	Percentage area	The outer surface area of the fruit not having spots/bruises/scratches and alike.
Color Features		
2	Red mean	Average of the red channel of the fruit image
3	Green mean	Average of the green channel of the fruit image
4	Blue mean	Average of the blue channel of the fruit image
5	Hue mean	Average of the hue channel of the fruit image. Original image is converted to HSV color model
6	Saturation mean	Average of the saturation channel of the fruit image. Original image is converted to HSV color model
7	Value mean	Average of the value channel of the fruit image. Original image is converted to HSV color model
8	Y mean	Average of the Y channel of the fruit image. Original image is converted to YCbCr color model
9	Cb mean	Average of the Cb channel of the fruit image. Original image is converted to YCbCr color model
10	Cr mean	Average of the Cr channel of the fruit image. Original image is converted to YCbCr color model
Gray level co-occurrence and texture features		
11	Contrast	A measure of the intensity contrast between a pixel and its neighbor over the entire image
12	Correlation	A measure of how correlated a pixel is to its neighbor over the entire image
13	Energy	Sum of squared elements in the GLCM
14	Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal
15	Entropy	A statistical measure of randomness, used to characterize the texture of an image

### Spatial domain features

Spatial domain can be defined as the section of the real plane spanned by the coordinates of an image (Gonzalez 2009). Image features are extracted by analyzing the representation of the image in spatial domain. In the present work, 15 spatial features are extracted. Each of the features is listed in Table 2.

### Wavelet features

Discrete wavelet transform (DWT) is expansively used in signal and image processing because of the fact that it captures image information pertaining to both spatial and frequency domains. On application of DWT, the image is decomposed into four sub-bands: (1) approximation sub-band, (2) horizontal detailed sub-band, (3) vertical detailed sub-band and (4) diagonal detailed sub-band. Further levels of decompositions are applied for the approximation sub-band. In the present study two-dimensional DWT is performed using three familiar wavelet families, namely, haar, daubuchies db2 and symlet (sym2). Wavelet transform for each wavelet family is done at three levels. Seven features, consisting of two statistical features and five gray level co-occurrence textural features, are computed at each level, resulting into a total of 252 wavelet features for each image. The mean and standard deviation are the statistical features. Five textural features are outlined in Table 2.

### Artificial neural network (ANN) training

A feed-forward back propagation network is used for training. Spatial features and wavelet features are trained separately. The feature dataset is randomly divided into three parts: (1) 60% for training, (2) 20% for validation and (3) 20% for testing. The network is trained for 20 iterations. Spatial features are trained with a neural network consisting of 15 input neurons, five hidden neurons and one output neuron. Wavelet features are trained with a neural network consisting of 252 input neurons, five hidden neurons and one output neuron. The transfer functions used are enumerated in Table 3.

### Support vector machines (SVM) training

Support vector machines are a set of associated supervised learning methods used for classification and regression. Given a set of training examples, each marked as belonging to one of the two categories, an SVM training algorithm develops a model

Table 3. Training functions used for training the artificial neural networks

Hidden layer	Hyperbolic tangent sigmoid transfer function
Output layer	Linear transfer function
Back propagation network training function	Levenberg–Marquardt backpropagation
Back propagation weight learning function	Gradient descent with momentum weight and bias learning function

that determines whether a new example falls into one category or the other. Support vector machines perform well in many machine learning applications. It searches for the linear hyper plane that splits the positive and negative samples with the largest margin. Depending on the training data, SVM elects an optimal hyperplane that can classify the given input either as healthy or diseased. In the present work, SVM has elected the training and testing set randomly. Linear kernel function is used with a scale of 1.5. The predictor data is standardized by setting it to ‘true’.

## RESULTS

Table 4 depicts the overall results and analysis. ANN and SVM training functions are run for 10 iterations. In each of the iteration, the accuracy of the classification is noted down. The average accuracy is then calculated. Standard deviation is also calculated.

From the results, following observations can be made:

1. When we employ spatial domain features for classifying pomegranate images, ANNs outperformed SVMs. This is due to the fact that ANNs present a deep architecture with hidden layers, bias and back propagation nature.
2. When we employ wavelet features for classifying pomegranate images, ANNs outperformed SVMs.
3. Using wavelet features for training showed an improved accuracy when compared to spatial domain features. And also, the standard deviation of all the populations in ANNs is less when compared to that of SVMs. This is due to the fact that wavelet transforms offers multi-resolution

analysis of the images. Wavelet transforms analyze the fine details of the image. But this is unlike in spatial analysis.

Moreover, the results of training are different for different iterations. This is due to the fact that in case of ANNs, the weights and biases are assigned randomly during each training session. And input data is also divided randomly for each iteration, taking random order of input/output pairs. In case of SVMs, random partitions are created for cross validation. Such inconsistencies can be avoided by using fixed weights and biases in case of ANNs and using fixed partitions in case of SVMs.

From the above observations, the overall result analysis can be depicted as shown in table 5. Hence it can be concluded that wavelet features can be used in the quality assessment process of the agriculture produce with artificial neural networks for improved accuracy.

We also would like to infer that the findings obtained by our method are highly promising in comparison to the works in the literature review.

Table 4. Result of analysis for sorting pomegranate fruits

Iterations	Accuracy using support vector machines (in %)		Accuracy using artificial neural networks (in %)	
	wavelet features	spatial features	wavelet features	spatial features
1	82.35	70.6	91.5	69.2
2	64.71	64.73	94	85.5
3	76.48	70.61	95.7	82.9
4	94.12	64.72	97.4	82.1
5	82.35	82.35	96.6	88
6	70.6	82.37	91.5	77.8
7	76.47	76.46	90.6	82.9
8	82.35	76.5	86.3	86.3
9	94.12	88.24	90.6	85.5
10	76.48	88.25	92.3	79.5
Average	80.00	76.48	92.65	81.97
Standard deviation	8.80	8.31	3.17	5.18

Table 5: Overall result analysis

Feature set	Support vector machines	Artificial neural networks
Spatial features	76.48	81.97
Wavelet features	80.00	92.65

## DISCUSSION

Automated qualification of fruits plays a vital role in post-harvest marketing. The present experiment addresses this issue in regard to pomegranate fruits with the help of a machine vision system. The experiment began with image acquisition using the image acquisition compartment. After pre-processing of the acquired images, spatial and wavelet features were extracted and analyzed with the help of two machine learning techniques: ANN and SVM.

On critical observation of the results given in the Tables, it has been found that wavelet features are well suited for the purpose stated above and ANN results are promising in comparison to the methods of the previous research works in this regard. The work by Ghazali et al. (2007) made use of DWT features and demonstrated an accuracy of 87.25% in classifying broad and narrow weeds. Comparatively, our investigation showed an improved accuracy of 92.65% with the use of DWT features. Most of the works, such as Teimouri et al. (2014), Font et al. (2014), Clement et al. (2013), Jamil et al. (2009), Omid et al. (2010), and Rocha et al. (2010) considered only spatial domain features for the purpose of qualification of the agricultural produce. However, there is a large scope for analyzing the images of the agricultural produce under multiresolution processing, which was lacking in the previous works. And our study has achieved it successfully with a higher and promising accuracy by using DWT. Also, our results are almost close to the results of application of DWT in medical imaging (Dua et al. 2012), which showed an accuracy of 93.33%. In the study of Rocha et al. (2010), even though a fusion of features and classifiers technique were used, the results gave an accuracy of 85%, which is far less compared to our investigation. The results reported by Jamil et al. (2009) were 73.3% accurate in grading palm oil FFB because of the use

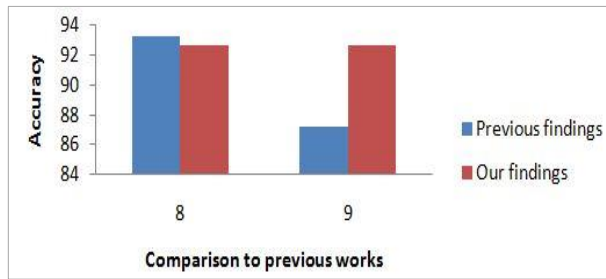


Fig. 4. Comparison to previous findings that considered DWT features

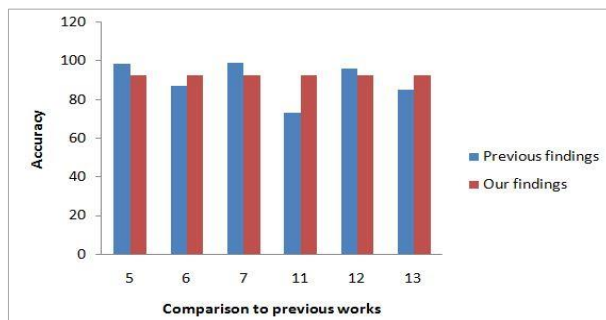


Fig. 5. Comparison to previous findings that did not consider DWT features

Table 6. Comparison with previous research works

Previous work	Accuracy (%)	DWT features considered	Comparison to our findings
Dua et al. (2012)	93.33	Yes	Almost equal results
Ghazali et al. (2007)	87.25	Yes	Improved accuracy
Teimouri et al. (2014)	98.82	No	Accuracy is less. But wavelet features are considered
Font et al. (2014)	87	No	Improved accuracy
Clement et al. (2013)	99	No	Accuracy is less. But wavelet features are considered
Jamil et al. (2009)	73.3	No	Improved accuracy
Omid et al. (2010)	96	No	Accuracy is less. But wavelet features are considered
Rocha et al. (2010)	85	No	Improved accuracy

of spatial features alone. This value is far less compared to 92.65% in our study. Since palm oil FFB has fruits and thorns together, there is a need to study such images under multiresolution. This requires the use of DWT so that higher classification rates can be obtained. Table 6 summarizes the discussions in regard to comparison of previous works and our findings. These comparisons are also depicted in the form of bar charts as shown in figures 4 and 5. Figure 4 shows the comparison of our results with the previous findings that considered DWT features and figure 5 shows the comparisons with studies that did not consider DWT features.

In addition, we obtained the classification rate of 81.97% considering spatial domain features. This result is almost similar to Font et al. (2014) and Rocha et al. (2010); however, the accuracy is higher when compared to Jamil et al (2009).

From the literature survey, we found that there was no significant research on the qualification of pomegranate fruits. Since, the post-harvest handling of pomegranates needs some of the problems to be addressed as discussed in the introduction section, there is a large scope for developing machine vision systems for sorting and grading of pomegranates. Since India is one of the largest exporters of pomegranates to the world, such automated systems become necessary in the current post-harvest markets. Hence our method is a remarkable beginning towards quality assurance of pomegranates in detail.

## CONCLUSION

In the present work, a machine vision system is developed to assess quality of pomegranate fruits. Initially images of two classes of pomegranate fruits are captured and dataset is built. An efficient pre-processing algorithm is applied on each of the image to tackle the presence of light during image capture. Two sets of features are then extracted for each of the image: spatial domain features and wavelet features. Artificial neural networks (ANNs) and support vector machines (SVMs) are then trained separately using both of the feature sets. Accuracies of results of the experiments are 76.483% using SVMs and spatial feature set, 80% using SVMs and wavelet features, 81.97% using ANNs and spatial features and 92.65% using ANNs and wavelet features.

Hence it can be concluded that ANNs outperform SVMs and wavelet features outperform spatial domain features.

However, there are some drawbacks of the present study. Future work need to be carried out to address them. Since the fruit samples collected consists of two cultivars, Bhagwa and Ganesh, there is a need for the analysis of other important cultivars. Since 252 wavelet transform features are extracted for each image, there is a scope to apply feature ranking and reduction techniques.

#### REFERENCES

- APEDA 2015. Pomegranate. In: Study on identification of export oriented integrated infrastructure for agri products from Maharashtra & Gujarat. Agriculture Produce Export Development Authority, pp. 20–22.
- Arefi A., Motlagh A.M., Mollazade K., Teimourlou R.F. 2011. Recognition and localization of ripen tomato based on machine vision. *Australian Journal of Crop Science* 5(10): 1144–1149.
- Babu K.D., Marathe R.A., Jadhav V.T. 2012. Post harvest management of pomegranate. ICAR – National Research Centre on Pomegranate, Solapur, India, 116 p.
- Benagi V.I., Nargund V., Balikai R., Ravikumar M. 2009. Pomegranate – Identification and Management of Diseases, Insect Pests and Disorders. University of Agricultural Sciences, Dharwad, India.
- Clement J., Novas N., Gazquez J.A., Manzano-Agugliaro F. 2013. An active contour computer algorithm for the classification of cucumbers. *Computers and Electronics in Agriculture* 92: 75–81. DOI: 10.1016/j.compag.2013.01.006.
- Deepa P., Geethalakshmi S.N. 2011. Improved watershed segmentation for apple fruit grading. *Proceedings of the International Conference on Process Automation, Control and Computing, IEEE*, 5 p. DOI: 10.1109/pacc.2011.5979003.
- Dua S., Acharya U.R., Chowriappa P., Sree S.V. 2012. Wavelet-based energy features for glaucomatous image classification. *IEEE Transactions on Information Technology in Biomedicine* 16(1): 80–87. DOI: 10.1109/titb.2011.2176540.
- Font D., Tresanchez M., Pallejà T., Teixidó M., Martínez D., Moreno J., Palacín J. 2014. An image processing method for in-line nectarine variety verification based on the comparison of skin feature histogram vectors. *Computers and Electronics in Agriculture* 102: 112–119. DOI: 10.1016/j.compag.2014.01.013.
- Ghazali K.H., Mansor M.F., Mustafa M.M., Hussain A. 2007. Feature extraction technique using discrete wavelet transform for image classification. *Proceedings of the 5<sup>th</sup> Student Conference on Research and Development, IEEE*, 4 p. DOI: 10.1109/scored.2007.4451366.
- Gonzalez R.C., Woods R.E., Eddins S.L. 2009. *Digital Image Processing Using MATLAB*, 2<sup>nd</sup> edition. Gatesmark Publishing, 827 p.
- Hazra T.K., Guhathakurta R. 2016. Comparing wavelet and wavelet packet image denoising using thresholding techniques. *International Journal of Science and Research* 5(6): 790–796. DOI: 10.21275/v5i6.nov164305.
- Jamil N., Mohamed A., Abdullah S. 2009. Automated grading of palm oil fresh fruit bunches (FFB) using neuro-fuzzy technique. *Proceedings of the International Conference of Soft Computing and Pattern Recognition, IEEE*, pp. 245–249. DOI: 10.1109/socpar.2009.57.
- Mustafa N.B.A., Ahmed S.K., Ali Z., Yit W.B., Abidin A.A.Z., Sharif Z.A.M. 2009. Agricultural produce sorting and grading using support vector machines and fuzzy logic. *Proceedings of the International Conference on Signal and Image Processing Applications, IEEE*, pp. 391–396. DOI: 10.1109/ic-sipa.2009.5478684.
- Omid M., Abbasgolipour M., Keyhani A., Mohtasebi S.S. 2010. Implementation of an efficient image processing algorithm for grading raisins. *International Journal of Signal and Image Processing* 1(1): 31–34.
- Rocha A., Hauagge D.C., Wainer J., Goldenstein S. 2010. Automatic fruit and vegetable classification from images. *Computers and Electronics in Agriculture* 70(1): 96–104. DOI: 10.1016/j.compag.2009.09.002.
- Teimouri N., Omid M., Mollazade K., Rajabipour A. 2014. A novel artificial neural networks assisted segmentation algorithm for discriminating almond nut and shell from background and shadow. *Computers and electronics in agriculture* 105: 34–43. DOI: 10.1016/j.compag.2014.04.008.
- Youwen T., Tianlai L., Yan N. 2008. The recognition of cucumber disease based on image processing and support vector machine. *Proceedings of the Congress on Image and Signal Processing* 2: 262–267. DOI: 10.1109/cisp.2008.29.