Feature Extraction in X-ray Images for Hazelnuts Classification

Ikramullah Khosa and Eros Pasero Department of Electronics and Telecommunications Politecnico di Torino Torino, Italy ikrmullah.khosa@polito.it eros.pasero@polito.it

Abstract—In the food industry, the importance of automatic detection and selection of raw food ingredients is increasing. In this paper, a method for real time automatic detection, segmentation and classification of hazelnuts using x-ray images is presented. Automatic extraction of independent nut images is made using image processing techniques. To extract meaningful features, moment invariants and texture properties are calculated on global level as well as from co-occurrence matrices. Principal component analysis is applied on features to achieve orthogonality in addition to dimensionality reduction. An anomaly detection algorithm is used for classification. Multivariate Gaussian distributions are calculated for model estimation using training data. Results are calculated on test data by using the threshold value obtained from best validation outcome. The classifier showed 98.6% correct classification rate for negative examples with 0% false negative rate.

I. INTRODUCTION

Autonomous quality assessment is gaining attention in food industry on product level as well as on ingredient level. One of the potential reasons is to develop an efficient and reliable autonomous production system to gain customer satisfaction about product quality using cutting-edge technology. Quality assessment of food products only at distribution line may lead to waste of resources, especially when raw ingredients are significantly smaller in size (like nuts) and need unusual attention before usage. In such a case, a sophisticated inspection system is required for ingredients before they proceed to the production line. Usually human visual inspection is carried out for initial selection of ingredients. In case of nuts, it is not applicable since internal inspection and detection of damage or infestation is not possible for a human without causing any damage to the ingredient. For the internal analysis, x-ray imaging has been employed in recent years in food and agriculture industry [1]-[7]. This is becoming a stateof-the-art technology for internal inspection of product at different levels of production. It provides the internal details of ingredients and reveals the presence of possible infestation inside without causing any damage to the ingredient itself. For a real time in-line inspection, it is essential to detect and segment each independent ingredient before making assessment about its quality. Previously, different image processing techniques have been employed for detection and segmentation of area of interest in the image [1], [8]. Then, it is

important to extract meaningful features from images which may enhance strong properties of images. Efforts have been made in extracting features from images for quality inspection using statistical, spectral, texture and geometric properties [3], [4], [9], [10]. Histogram features were used for classification of nuts [3], [11]. Texture properties like contrast, energy, entropy, homogeneity, uniformity and diagonal moment were employed for content based image classification [12]. Spectral imaging was used for quality inspection of cherries [13]. Texture features are more often used for quality inspection in these studies. In contrast, moment invariant [22] have not used often for quality assessment of food ingredient. These have been mostly employed in classification problems related to medical applications [20], [21]. However these features have potential to contribute in quality assessment problems since these are insensitive to translation, scaling, mirroring and rotation. Considering the food ingredients for quality assessment, x-ray images of pistachio nuts have been mostly used in aforementioned studies but hazelnuts in particular have examined less often.

Moreover, selection of an appropriate classifier is essential for an optimized solution. A classifier may fail with excellent features due to selection of in-appropriate classification technique. Many classification methods have been employed in classification of food and agricultural products including statistical discriminant classifiers, k-nearest neighbor classifier and Artificial Neural Networks (ANN) [1], [2], [7], [13], [14]. In particular, ANN have been extensively used in such classification problems. They have been proven as strong candidates for a binary or multiclass classification task where sufficient amount of labeled data for all classes are available. In contrast, anomaly detection [15]–[17] has not been employed more often in nuts classification. It is useful for a classification problem where a large percentage of data belong to one particular class and a small percentage to the other.

The aim of the our study was:

- to develop a method for detection and extraction of each individual hazelnut image from captured x-ray image of collective nuts in real time;
- to calculate meaningful features which hold information appropriate for quality assessment;



Fig. 1: Proposed setup for real time hazelnuts classification

 to select a good classification algorithm to produce optimum results.

The paper is organized in six sections.

Nutmeat extraction (extraction of part of the interest from the image) is explained in section 2. Section 3 illustrates the feature extraction. Classifier's details are included in section 4. Performance evaluation is presented in section 5. Conclusions are shown in section 6.



(a) Sample sub-image



(b) Binary image obtained after thresholding



(c) True nutmeat regions with centroids and bounding boxes

Fig. 2: Individual nutmeat detection with corresponding centroids and bounding boxes

II. NUTMEAT EXTRACTION

Considering a real time scenario, an x-ray scanner is supposed to be attached above the assembly line taking x-ray images of belt with raw hazelnuts laid in a serial manner. The picture of proposed setup for a real time hazelnuts classification is shown in Fig. 1. The x-ray images are stored, nuts are individually identified, feature extraction and classification is carried out, and decision vector is passed to the decision belt. The appropriate mechanical actuator is operated and unhealthy nuts are separated from the healthy ones.

To extract an individual nut image, we developed a method including detection, segmentation and extraction operations using image processing techniques. A sample of X-ray subimage is shown in Fig. 2a. Initially, each hazelnut is detected independently. The x-ray images of hazelnuts are obtained in RGB color space. For the detection of independent nutmeat, we initially converted it to a gray scale image [18]. Later, global thresholding is applied to detect the area of interest in the image [18]. At this step, many regions in the image are detected including real nutmeat regions and few external objects as shown in Fig. 2b. To decide whether a region is a real nutmeat region, the areas of all detected regions are calculated. After estimating the area of a normal hazelnut, we thresholded all the regions with T=A \pm 15%, where A is the area of a normal size hazelnut. Now, it is needed to extract the nutmeat from the original image. One choice is to multiply the detected region with the original image and retrieve the nutmeat region from the image. A sample nutmeat region extracted by using this method is shown in the Fig 2a. In this case, some information on the image boundary is lost. To overcome this, we changed the threshold value. Consequently, more boundary information is retained. But at the the same time, many background pixels are included in the extracted image as shown in Fig. 3b. So, a complete extraction of only the foreground nutmeat region is not possible by simple

threshloding. In addition, one fixed threshold can not insure the similar thresholding result on all nut images, since there is a high probability of variation in range of pixel intensities in different nut images. Hence, extraction of nutmeat only by thresholding is not a solution. To achieve a neat and complete extraction of a nutmeat region, we calculated centroids (coordinates of central pixel) of each region and bounding boxes (dimension of rectangular area) around each true nutmeat region detected (see Fig. 2c) [18]. Bounding boxes are calculated using boundary pixels in the horizontal and vertical axes of binary image (see Fig 3b) obtained after thresholding. Later, we extracted each nutmeat independently from original x-ray image by using its corresponding bounding box. A nutmeat sample extracted using its centroid and bounding box with padding is shown in the Fig. 3c. The information loss in case of extraction by simple thresholding and multiplying with original image (using different threshold values) in comparison to extraction using bounding box is shown in terms of histograms in Fig. 3d, 3e and 3f.

III. FEATURE EXTRACTION

Texture features have been mostly used for quality assessment of nuts. Since in x-ray imaging, nuts appear with variable size and at different orientations in the image, so we considered moment invariant features which are insensitive to rotation, scaling and translation. We calculated texture properties from images on the global level as well as from co-occurrence matrices.

A. Global Texture Features

We extracted six feature on the global level from each of the sample image. For an image I with highest intensity level N, these properties are calculated as;

1) Mean: It is the measure of average intensity

$$\mu = \sum_{i=0}^{N-1} u_i p(u_i)$$
 (1)



(a) Lower value of thresh-(b) Larger value of thresh-(c) Bounding box dimenold old sions



Fig. 3: Extracted samples by using different methods and their respective histograms

where u_i represents the *i*th intensity level and $p(u_i)$ is its probability.

2) Standard Deviation: It is the measure of average contrast

$$\sigma = \sqrt{\sum_{i=0}^{N-1} (u_i - \mu)^2 p(u_i)}$$
(2)

3) Smoothness: It is calculated between 0 and 1, higher the excursions in the intensity level, higher will be the value of smoothness

$$S = 1 - \frac{1}{1 + \sigma^2}$$
(3)

where σ^2 is variance.

4) Third Moment: This is the measurement of skewness of a histogram

$$M_3 = \sum_{i=0}^{N-1} (u_i - \mu)^3 p(u_i)$$
(4)

5) Uniformity: It is a measure opposite to smoothness, a constant intensity image provides zero smoothness while maximum uniformity

$$U = \sum_{i=0}^{N-1} (p^2(u_i))$$
(5)

6) Entropy: It measures the randomness in the image

$$E = -\sum_{i=0}^{N-1} p(u_i) \log_2 p(u_i)$$
(6)

B. Texture Feature from Co-occurrence Matrix

A Co-occurrence matrix is used for texture analysis of an image and provides the information about how often specific combinations of gray levels occur in an image [19]. Its advantage over histogram is that it provides the frequency of pair of pixel intensities and is sensitive to rotation or translation, while histogram gives only the frequency of pixel intensities and is insensitive to rotation or translation. Gray Level Co-occurrence Matrix (GLCM) is calculated according to a position operator (also called offset) which defines the distance and the angle parameter. We have calculated four GLCMs from each of the sample images with a distance of one pixel at angles of 0, 45, 90 and 135 respectively. All intensity levels present in the image are considered for the calculation of GLCMs. The position operator used for calculation of cooccurrence matrices, with distance and angles is shown in the Fig. 5. A GLCM for an image I of size $m \times n$ is calculated as

$$C_{\Delta x,\Delta y}(i,j) = \sum_{p=1}^{m} \sum_{q=1}^{n} \begin{cases} 1, \text{ if } I(p,q) = i \\ \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0, \text{ otherwise} \end{cases}$$
(7)

where Δx and Δy represent horizontal and vertical distances respectively. After computing GLCMs, we have calculated the following four texture properties from each of them and obtained a total of 16 texture features from GLCMs.



(a) Healthy



(b) Damaged



(c) Infected

Fig. 4: Image samples of healthy and unhealthy hazelnuts



Fig. 5: Position operator with distance and angles

1) Contrast: It provides the intensity contrast between any pixel and its neighbor over the entire image

$$C = \sum_{i,j} (i-j)^2 c_{ij} \tag{8}$$

where i and j are pixel intensities, c_{ij} is their count of cooccurrences according to the specified position operator.

2) Correlation: It is a statistical measure of correlation between any pixel and its neighbor

$$c = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)c_{ij}}{\sigma_i \sigma_j}$$
(9)

where μ_i and μ_j are mean values, σ_i and σ_j are standard deviations of pixel intensities in rows and columns in the respective GLCM.

3) Energy: This is equal to sum of squared elements of the GLCM

$$e = \sum_{i,j} c_{ij}^2 \tag{10}$$

4) Homogeneity: It is measured between 0 and 1 and defines how close is the distribution of GLCM elements to the GLCM diagonal

$$H = \sum_{i,j} \frac{c_{ij}}{1 + |i - j|}$$
(11)

C. Moment Invariant Features

We calculated a set of moment invariant features. A two dimensional moment of order p + q of a an image I(x, y) is calculated

$$m_{pq} = \sum_{x} \sum_{y} x^{p} y^{q} I(x, y) \qquad p, q = 0, 1, 2, 3, \dots$$
 (12)

The corresponding central moment is defined as

$$\mu_{pq} = \sum_{x} \sum_{y} (x - \bar{x})^p (y - \bar{y})^q I(x, y)$$
(13)

where $\bar{x} = \frac{m_{10}}{m_{00}}$ and $\bar{y} = \frac{m_{01}}{m_{00}}$ The normalized central moment of order p + q is defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \qquad p, q = 0, 1, 2, 3, \dots$$
(14)

where $\gamma = \frac{p+q}{2} + 1$ p+q = 2, 3, ...A set of seven 2-D moment invariants which are insensitive to translation, scaling and rotation are derived and calculated as

$$M_{1} = \eta_{20} + \eta_{02}$$

$$M_{2} = (\eta_{20} - \eta_{02})^{2} + 4\eta_{11}^{2}$$

$$M_{3} = (\eta_{30} - 3\eta_{12})^{2} + (3\eta_{21} - \eta_{03})^{2}$$

$$M_{4} = (\eta_{30} + \eta_{12})^{2} + (\eta_{21} + \eta_{03})^{2}$$

$$M_{5} = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^{2} - (3(\eta_{21} + \eta_{03})^{2}] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$

$$[(3\eta_{30} + \eta_{12})^{2} - (\eta_{21} + \eta_{03})^{2}]$$

$$M_{6} = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^{2} - (\eta_{21} + \eta_{03})^{2}] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$M_{7} = (3\eta_{21} - 3\eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^{2} - 3(\eta_{21} + \eta_{03})^{2}] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\[(3\eta_{30} + \eta_{12})^{2} - (\eta_{21} + \eta_{03})^{2}]$$

D. Dimensionality Reduction

We have calculated a total of 29 features including 6 textures features on the global level, 16 features from cooccurrence matrices and 7 moment invariant. To increase the computational efficiency of the classifier, we applied Principal Component Analysis (PCA) on extracted features [23], [24]. PCA is a technique used to convert a large feature set into a small feature set providing orthogonality among features. Principal components are obtained from original features (by orthogonal transformation) and arranged in such a way that first component contains the highest variance and so on. We applied PCA to already extracted feature set. Later, we selected first three principal components to produce the new feature set. we obtained a new set with three features which contain 99% variance of original features. Then we used this new feature set, consisting of 3 features, for classification purposes.

IV. CLASSIFIER SELECTION

The x-ray image database contains 748 images of healthy hazelnuts, 20 images of damaged nuts and 20 of infected hazelnut. 95% of examples belong to "good" category only 5% examples to "bad" category. For this data, we selected anomaly detection algorithm for the detection of bad nuts. This is a technique which is used to detect an unexpected outcome. We employed this technique in a semi-supervised way, for the detection of abnormal examples (unhealthy nuts) among the the normal examples (healthy nuts).

A. Anomaly Detection Algorithm Development

Images of infected nuts and damaged nuts are combined and labeled as positive examples while images of healthy nut samples are labeled as negative examples. We have randomly selected 600 image samples of healthy nuts as training data and used them for model building. The mean and variances of these features are calculated. The objective is to set up a model using normal examples which statistically defines a range for an unseen example to be a normal one. The one outside the range is flagged as abnormal (or anomaly). A multivariate Gaussian distribution is calculated from training data as

$$P = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$
(16)

where μ and Σ are mean and covariance matrix respectively. The multivariate normal distribution model is built using normal examples. The rest of the data is divided as; *Cross Validation Data* = 74 negative examples + 20 positive examples

 $Test \ Data = 74$ negative examples + 20 positive examples.

V. PERFORMANCE EVALUATION

The model is tested initially on cross validation data [26]. By using the mean and covariance matrix, calculated from training data, probability density function for the cross validation data P_{val} is calculated with the help of the expression given in (16). Then, we quantized P_{val} by a step size defined as;

$$Step \ size = \frac{max(P_{val}) - min(P_{val})}{600} \tag{17}$$

Each of these 600 values are then used as threshold to evaluate the validation data. The objective is to estimate a threshold ϵ , provided by best cross validation outcome. For each value of ϵ , number of outliers in the validation data are calculated as

$$Outlier = \begin{cases} 1, \text{if } P_{val} < \epsilon \\ 0, \text{ otherwise} \end{cases}$$
(18)



Fig. 6: Probabilities of test samples

The classifier's performance is evaluated using validation data and F_1 score is calculated for each value of threshold. The threshold with highest F_1 score is taken and used to classify the test data.

Statistics are calculated including parameters: precision, recall and F_1 score including number of outliers for each value of ϵ . These parameters are defined as

$$\begin{aligned} Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F_1 \, Score &= 2 \times \frac{Precision \times Recall}{Precision + Recall} \end{aligned}$$

where

TP = Total number of unhealthy nuts classified correctly as positive examples.

FP = Total number of healthy nuts classified incorrectly as positive examples.

FN = Total number of unhealthy nuts classified incorrectly as negative examples.

 F_1 score is an effective and conveniently single parameter to evaluate the classifier's performance. It ranges between 0 and 1.

Once test set results are calculated, the data are randomized. All three sets: training, validation and test set are randomly selected again. Features extraction, model building, testing and optimizing the model on validation data, and calculation of test set results is carried out all over again. The process is repeated seven times and an average of test and validation set result is summarized in Tab. I.

A. Results and Discussion

The classifier produced fairly good average results for both the categories. The key task for this classification problem is to identify each bad nut (damaged or infected) correctly, while achieving a maximum possible correct classification rate for good nuts. The results show that maximum possible anomalies (unhealthy nuts) are detected correctly (see Tab. I).

TABLE I: Classification Results for Cross Validation and Test Data

Cross Validation Data			Test Data							
No of Outliers	Best Threshold	Best F_1 Score	No of Outliers	TP	FN	TN	FP	Precision	Recall	F_1 Score
21	$1.77e^{-8}$	0.9756	21	20	0	73	1	0.9524	1	0.9756

Figure 6 shows the lowest probabilities for tail end unhealthy samples presented to the classifier, detected as anomalies. Also the classifier showed good capability of correct recognition of normal examples (healthy nuts). An average of optimized results obtained from seven rotated validation data sets is also presented in Tab. I. Since the data are rotated and model is optimized every time with different samples, the results also empower the robustness of the classifier with calculated features. The Receive Operative Characteristics curve (ROC) [27] is calculated for the test data and presented in Fig. 7. Results can be compared with previous works in similar classification applications. Keagy in [11] used different set of features, extracted from x-ray image of pistachio nuts and achieved 95.8% test accuracy for good nuts while 81.9% for bad nuts using 14 quadratic MRDF features. The database however was quiet larger (9 times) than the one we used here. Casasent used histogram features in [9] and statistical features in [14] for pistachio nuts x-ray image classification, and achieved 89% and 88% of test set accuracy respectively. Motto Ros in [25] used fuzzy features for defect detection in pistachio nuts and achieved 99.6% correct recognition with 0.3% of false positive rate. Considering the fact that the size of database and even the data used in these works were different, so a fair comparison is not possible. However, the obtained results of 0% false positive rate and 98.6% true negative rate using this database are quiet encouraging and competitive.

VI. CONCLUSION

In this paper, we developed a method for a possible real time detection, segmentation and classification of hazelnuts. Image processing techniques were employed to develop an extraction mechanism of true nutmeat from the collective image. The features were moment invariant and texture properties: on the global level, and from co-occurrence matrices. Gray Level Co-occurrence Matrices were calculated from each sample nut image at angles of 0,45,90 and 135 with one pixel depth. Principal Component Analysis was applied to obtain a smaller feature space. Anomaly detection algorithm was used to detect the unhealthy nuts among the healthy ones. The model was built using training data containing negative examples only. The threshold was selected from optimized validation outcome and used to classify the test data. The classifier showed 100% true positive rate while achieving 98.6 % true negative rate.

ACKNOWLEDGMENTS

We would like to thank Ferrero and Soremartec for their long-standing support of our research activity. We also thank Dr. A. Boscolo and Dr. L. Placentino, for providing us with the set of x-ray images used in this work. This study was funded by ITACA, a project financed by the European Union,



Fig. 7: The ROC space for test data

the Italian Ministry of Economy and Finance and the Piedmont Region.

REFERENCES

- D.S. Narvankar, C.B. Singh, D.S. Jayas and N.D.G. White, "Assessment of soft X-ray imaging for detection of fungal infection in wheat", *Biosystems Engineering*, Volume 103, Issue 1, May 2009, Pages 49-56, ISSN 1537-5110.
- [2] S. Neethirajan, D.S. Jayas, N.D.G. White, "Detection of sprouted wheat kernels using soft X-ray image analysis", *Journal of Food Engineering*, Volume 81, Issue 3, August 2007, Pages 509-513, ISSN 0260-8774.
- [3] P.M. Keagy, B. Parvin and T.F. Schatzki, "Machine recognition of navel orange worm damage in X-ray images of pistachio nuts" *Lebensmittel-Wissenschaft+ Technologie* 29.1-2 (1996): 140-145.
- [4] T.F. Schatzki, et al. "Defect detection in apples by means of X-ray imaging" *Transactions of the ASAE* 40.5 (1997): 1407-1415.
- [5] P.M. Keagy and T.F. Schatzki, "Machine recognition of weevil damage in wheat radiographs" *Cereal chemistry* 70 (1993): 696-696.
- [6] P.M. Keagy, B. Parvin, T. Schatzki, L.C. Lee, D. Casasent and D. Weber, "Expanded image database of pistachio x-ray images and classification by conventional methods", *Proc. SPIE*, volume 2907, pages 196-204, Nov. 1996.
- [7] A. Talukder, D. Casasent, H. Lee, P.M. Keagy and T.F. Schatzki, "New feature extraction method for classification of agricultural products from x-ray images", *Proceedings of SPIE-The International Society of Optical Engineering*, Vol. 3543, 1999.
- [8] D.P. Casasent, A. Talukder, W. Cox, H.T. Chang and D. Weber, "Detection and segmentation of multiple touching product inspection items", *Proc. SPIE 2907, Optics in Agriculture, Forestry and Biological Processing II*, pp.205-216, December 1996.
- [9] A. Talukder, D. Casasent, H. Lee, P. M. Keagy and T.F. Schatzki, "New feature extraction method for classification of agricultural products from x-ray images" *Proceedings of SPIE- The International Society for Optical Engineering* Vol. 3543, 1999.
- [10] A. Mueen, M.S. Baba and R. Zainuddin, "Multilevel feature extraction and X-ray image classification" *Applied Sciences* 7.8 (2007): 1224-1229.

- [11] P.M. Keagy, T.F. Schatzki, L.C. Le, D.P. Casasent and D. Weber, "Expanded image database of pistachio x-ray images and classification by conventional methods" *Proc. SPIE* Vol. 2907. 1996.
- [12] S.B. Park, J.W. Lee and S.K. Kim, "Content-based image classification using a neural network" *Pattern Recognition Letters* 25.3 (2004): 287-300.
- [13] D. Guyer and X. Yang, "Use of genetic artificial neural networks and spectral imaging for defect detection on cherries" *Computers and Electronics in Agriculture* 29.3 (2000): 179-194.
- [14] D.P. Casasent, M.A. Sipe, T.F. Schatzki, P.M. Keagy and L.C. Le, "Neural net classification of x-ray pistachio nut data" *Lebensmittel-Wissenschaft+ Technologie* 31.2 (1998): 122-128.
- [15] V. Chandola, A. Banerjee and V. Kumar, "Anomaly detection: A survey", ACM Comput. Surv. 41, 3, Article 15 (July 2009), 58 pages.
- [16] Z. Niu, S. Shi, J. Sun and X. He, "A survey of outlier detection methodologies and their applications", *Proceedings of the Third international conference on Artificial intelligence and computational intelligence*, September 24-25, 2011, Taiyuan, China.
- [17] A. Weiler, M.H. Scholl, F. Wanner and C. Rohrdantz, "Event identification for local areas using social media streaming data", *Proceedings of the ACM SIGMOD Workshop on Databases and Social Networks*, p.1-6, June 22-27, 2013, New York.
- [18] R.C. Gonzalez, R.E. Woods and S.L. Eddins, *Digital image processing using MATLAB* Vol. 2. Knoxville: Gatesmark Publishing, 2009.
- [19] R.M. Haralick, K. Shanmugam and I.H. Dinstein, "Texture features for image classification", *IEEE Transactions on Systems, Man and Cybernetics, SMC-3*, Issue 6, pp-610-621, November, 1973.
 [20] M. Pan, J. Tang and X. Yang, "An algorithm for medical image tilt
- [20] M. Pan, J. Tang and X. Yang, "An algorithm for medical image tilt correction using b-spline and moment invariant", *ICIC Express Letters*, vol.4, no.1, pp.57-64, 2010.
- [21] H. Pourghassem and H. Ghassemian, "Content-based medical image classification using a new hierarchical merging scheme", *Computerized Medical Imaging and Graphics*, vol.32, no.8, pp.651-661, 2008. Visual Pattern Recognition by Moment Invariants
- [22] Ming-Kuei Hu, "Visual Pattern Recognition by Moment Invariants", *IRE Transactions on Information Theory*, vol 8, issue 2, pp 179-187, February 1962.
- [23] H. Abdi and L.J. Williams, "Principal component analysis" Wiley Interdisciplinary Reviews: Computational Statistics 2.4 (2010): 433-459.
- [24] I. Jolliffe, "Principal component analysis" John Wiley & Sons, Ltd, 2005.
- [25] P. Motto Ros and E. Pasero, "Defects Detection in Pistachio Nuts Using Artificial Neural Networks", *Neural Nets and Surroundings*, Springer Berlin Heidelberg, 2013. 147-156.
- [26] R. Kohavi, "A study of cross validation and bootstrap for accuracy estimation and model selection", *Proceedings of the 14th IJCAI*, 1995, Montreal, Canada.
- [27] T. Fawcett, "An introduction to ROC analyses", *Pattern Recognition Letters*, vol.27, pp. 861-874, 2006.