Classification of Wheat Cultivars Using Image Processing and Artificial Neural Networks

Alireza Khoshroo
Yasouj University

Arman Arefi

Asad Masoumiasl
Yasouj University

Gholam-Hossein Jowkar
Zurich University of Applied Sciences

AGRICULTURAL COMMUNICATIONS, 2014, 2(1): 17-22.

Classification of Wheat Cultivars Using Image Processing and Artificial Neural Networks.

ALIREZA KHOSHROO[1*], ARMAN AREFI[2], ASAD MASOUMIASL[3] AND GHOLAM-HOSSEIN JOWKAR[4]
[1] Department of Agricultural Engineering, University of Yasouj, Yasouj, Iran.
[2] Young Researchers and Elite Club, Islamic Azad University, Sanandaj Branch, Sanandaj, Sanandaj, Iran.
[3] Department of Agronomy and Plant Breeding, University of Yasouj, Yasouj, Iran.
[4] Department of Computer Science and Engineering, Shiraz University, Shiraz, Iran.
*Corresponding Author: khoshroo@yu.ac.ir *(Accepted: 7 Dec. 2013)*

ABSTRACT
Classification of wheat grains has significant importance in determining the market value of wheat varieties. Wheat class identification is also necessary for plant breeders to predict yield and quality. In this study, classification of four Iranian wheat cultivars was carried out using morphological features and artificial neural networks. After preparing samples, 164 images of grains were acquired for each cultivar in a lighting chamber. Ten morphological features were extracted from images using image processing techniques. For classifying wheat varieties, various topologies of artificial neural networks (ANN) with different number of neurons in the hidden layers were developed. The nine important morphological features extracted from images were used as input for developed ANN. 60% of all samples were used for training networks. Validation of the developed ANN structures was done by 15% of samples while 25% of samples were used for evaluation of the networks. The best topology for ANN was 9-26-4. Results showed that overall classification accuracy of 85.72% was obtained for classification of wheat cultivars.

Keywords: ANN, cereal, data mining, machine vision, morphological features, seed morphology.

Abbreviations:
ANN: artificial neural networks; **MLP:** multi-layer perception.

INTRODUCTION

Wheat is one of the most important cereals because of its valuable nutrients. It is a major source of energy, protein, and dietary fibber in human nutrition (Mansing, 2010). Wheat provides approximately one-fifth of the total calorific input of the World's population (FAO, 2011). As a result, annual production of wheat has greatly increased in the last years. In terms of total production tonnages, the world production of wheat in 2011 was 701.39 million tons (FAO, 2011). In 2011, Iran with 13.5 million tons of wheat production has been ranked 14th in the world that shows increasing trend comparing with previous years (FAO, 2011).

Determination of wheat varieties is a necessary step for growers, processors and consumers. Classification of wheat grains has significant importance in determining the market value of wheat varieties. Quality of final products depends on the use of specific wheat varieties. Wheat class identification is also necessary for plant breeders to predict yield and quality (Shouche *et al*., 2001).

The morphological features of grains are heritable characters and play an important role in wheat variety recognition. In the current grain grading systems, grain type and quality are inspected through visual assessment by experts and trained inspectors which is subjective, and tedious. Computer vision system offers an objective and quantitative method for estimation of morphological parameters and quality of agricultural products to obtain quick and more reliable results (Arefi *et al*., 2011; Choudhary *et al*., 2008; Khoshroo *et al*., 2006).

Image processing based on morphology, colour and textural features of grains is necessary for different applications in the grain industry including assessing grain quality and discrimination of wheat classes. Various grading systems have been reported in literatures which use different morphological features for the classification of different cereal grains and cultivars (Shouche *et al*., 2001; Dubey *et al*., 2006; Zapotoczny *et al*., 2008; Masoumiasl *et al*, 2013). Utku (2000) developed a system to identify 31 bread wheat and 14 durum wheat cultivars using CCD video camera. The use of orthonormal transformation discriminated the bread wheat and durum wheat well, but the classification accuracy of cultivars within group was not satisfactory. Majumdar and Jayas (2000 a,b,c,d) applied digital image processing and discriminate analysis to develop a system for identification of different grain species. They used morphological, colour, textural and combination of these features to describe physical properties of the kernels.

In grain classification process, several techniques such as statistical, artificial neural networks and fuzzy logic have been used. Classification performances of different neural network topology were compared by using morphological features of Canada Western Amber Durum (CWAD) wheat, Canada Western Red Spring (CWRS) wheat, oats, rye and barley (Visen *et al*., 2004). The objectives of this study were to determine the utility of morphological features for classification of individual kernels of four Iranian rainfed wheat varieties and to find the best method for classifying the kernels of wheat with the lowest error of classification. So, combination of machine vision system and artificial neural network classifier were developed for discriminating between rainfed wheat grain cultivars.

MATERIALS AND METHODS

Overall six steps including sample preparation, image acquisition, pre-processing operations, feature extraction, feature selection and classification were performed to classify wheat cultivars (Fig. 1). Wheat cultivars classification steps are explained in following:

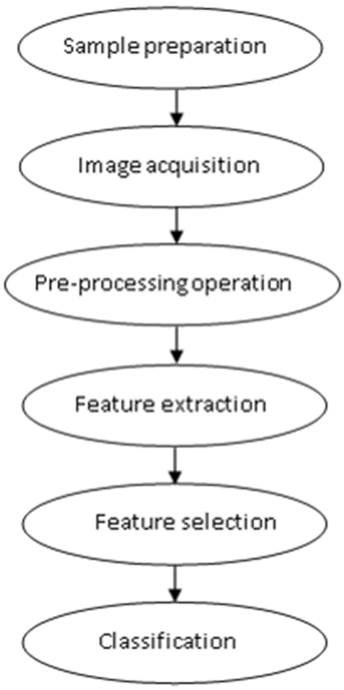


Fig. 1. Flowchart of wheat cultivars calassification process.

Sample Preparation:

In this study, identification of four Iranian wheat cultivars was carried out. 'Gohar', 'Dehdasht', 'Koohdasht 'and 'Seimareh' are wheat cultivars widely cultivated in the southern and south western region of Iran (Fig. 2). More than 1kg from each cultivar was obtained as a sample. The grains were cleaned manually to remove non-grain matter and damaged grains.

Image Acquisition:

An image capturing system was designed to take some standard pictures from the samples. The size of the chamber was 40×40×40 cm. Wheat samples were placed on the holder. In order to eliminate the shadow, the holder was covered with black material. Samples were illuminated using two parallel lamps (with one fluorescent tube in each lamp, model 391 Deluxe, Natural Daylight, 10W, Farhad Lighting Co., Tehran, Iran), which were equipped with light diffuser and electronic ballast. The two fluorescent tubes (391 mm) were placed 35 cm above and parallel to the sample holder. Moreover, wheat grains were separated manually to eliminate grain contact. A colour CCD camera (CNB, 560 TV line, model GA4162PF, Korea) was positioned horizontally in centre of the chamber and vertically over the sample holder at a distance of 40 cm. 164 colour images of grains of each cultivar in the RGB colour model were acquired. The images were sent via a TV capture device (Axtrom, XT-TV100, Korea) to a personal computer.

Pre-Processing and Segmentation Operations:

The first step in each image processing operation is known as pre-processing and segmentation operations. Pre-processing and segmentation operations guaranty the quality of the final result of analysis (Cheng *et al*., 2001). To modify the non-uniform distribution of background light intensity and to remove any external noise from the image background pre-processing and segmentation is necessary. Image pre-processing and segmentation was performed in MATLAB software (Version R2012a, the MathWorks Inc., MA, USA), as described below:

a) Obtaining grey images from the RGB space channels.
b) Obtaining binary image of samples using defined threshold values for R channels. In binary images, grains and background have the value equal with 1 and 0, respectively.
c) Removing noise using eroding by reconstruction operation.
d) Filling of holes in the binary image to obtain an actual binary image.
e) Multiplying of obtained binary images in R, G, and B channels.

f) Acquiring RGB images by combination of grey images obtained from the previous step.

The result of these steps is shown in Fig. 3.

Feature Extraction:

Extraction of robust features can be an important issue in the ultimate success of developed algorithm. There are different features for describing an object. Morphology is known as one of the most effective features that can be useful in discriminating various objects. The word morphology denotes visual and shape characteristics of an object. On this basis, Recognition of wheat cultivars was carried out based on the morphological parameters. Using image processing toolbox of MATLAB software (Version R2012a, The MathWorks Inc., MA, USA), ten morphological features were extracted for each kernel. The calculated features are perimeter, area, major axis length and minor axis length. The area of a region was defined as the number of pixels contained within its boundary and its perimeter was the length of its boundary. Correspondingly six morphological features were calculated as described below:

$$\text{Compactness}=\frac{4\pi(\text{area})}{(\text{perimeter})^2}$$

$$\text{Circularity}=\frac{(\text{perimeter})^2}{\text{area}}$$

$$\text{Area ratio}=\frac{\text{area}}{(\text{major axis})(\text{minor axis})}$$

$$\text{Aspect ratio}=\frac{\text{major axis}}{\text{minor axis}}$$

$$\text{Eccentricity}=\frac{2\sqrt{\text{major axis})^2-(\text{minor axis})^2}}{(\text{major axis})(\text{minor axis})}$$

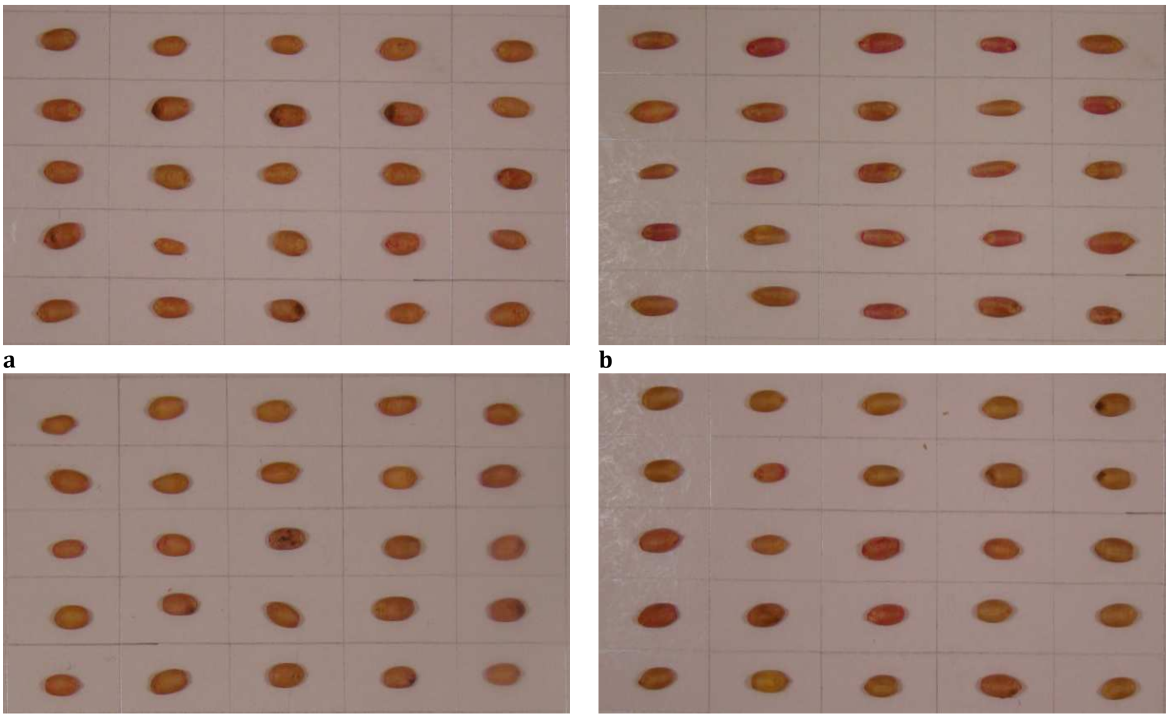
$$\text{Solidity}=\frac{\text{area}}{\text{convex area}}$$

Feature Selection:

After the morphological features were obtained for each image, feature selection was conducted to reduce the redundancy in the morphological feature set. Stepwise discriminant analysis was used to select a subset of the features for discriminating among the classes. The method is suitable for selecting the most effective morphological feature. Variables are chosen to enter or leave the model depending on the significance level of an F test from an analysis of covariance, where the variables already chosen act as covariates and the variable under consideration is the dependent variable

Intelligent Classification:

Classifying Four Iranian wheat cultivars was done using artificial neural network (ANN) based on morphological features using Neuro Solutions software. A multi-layer perception (MLP) network which is commonly used to classify objects was designed. MLPs often have one or more hidden layers of linear or non-linear neurons followed by an output layer.



a **b** **c** **d**

Fig. 2. Samples of wheat cultivars: a: 'Gohar', b: 'Dehdasht', c: 'Koohdasht', d: 'Seimareh'.

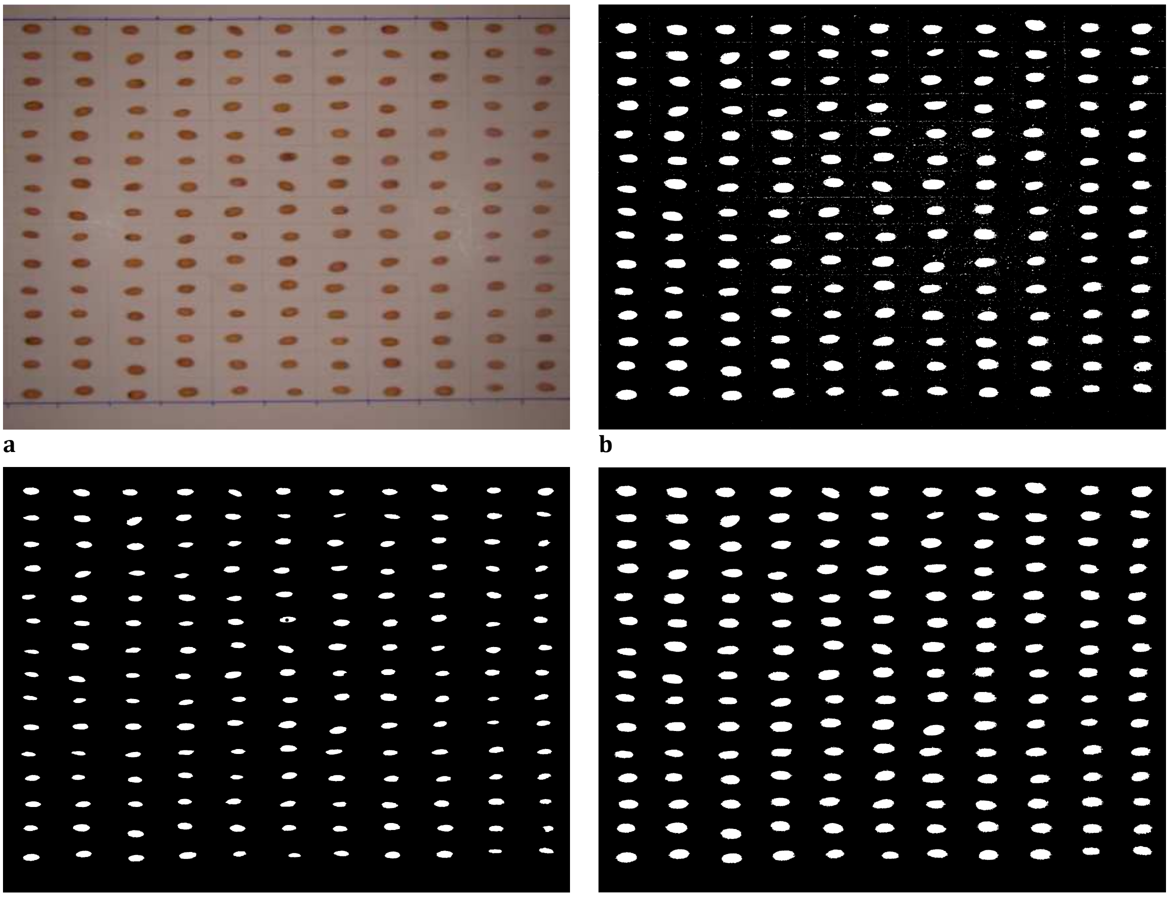


Fig. 3. Pre-processing image operations: a) original image, b) binary image resulted of red and blue subtraction, c) removing noises by erosion operation, d) reconstructing image and filling holes.

Several layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. Based on trial and error method, networks with one or two hidden layers with different number of neurons for hidden layers were developed. Number of neurons in output layer was four and the number of neurons in input layer was equal to number of input features. In the developed MLP models, the tangent sigmoid which is a non-linear transfer function was used for hidden layers. 60% of all samples were used for training the network to adjust the network weights. 15% of samples were used for validating the structure of network and remaining samples (25%) were used for evaluating the network in the testing step. In total, the training was performed 3 times and continued to minimize the mean squared error (MSE) between targets and outputs.

RESULTS AND DISCUSSIONS

Morphological features were computed from individual kernel images using image processing techniques. Some of morphological features were highly correlated with another and they did not contribute significantly to the morphology model. Table 1 shows the features in descending order of their contribution level to the classifier.

Table 1. Selection of the most effective features using stepwise discriminant analysis.

Features	ASCC†	Partial r^2
Aspect ratio	0.229	0.68
Compactness	0.432	0.72
Circularity	0.522	0.78
Eccentricity	0.539	0.20
Perimeter	0.556	0.07
Minor Axis	0.564	0.08
Solidity	0.568	0.04
Major Axis	0.573	0.02
Area ratio	0.575	0.01

† Average squared canonical correlation.

The Area feature was not detected as a robust feature and was removed automatically. This means that area is not able to discriminate between wheat varieties. It does not only decrease ANN performance but also increase processing time and ANN complexity. On the other hand, the aspect ratio (ASCC=0.23), Compactness (ASCC=0.43) and Circularity (ASCC=0.52) were the most important morphological features, respectively. It can be expected these features resulted in acceptable classification accuracy because they presented better description of different wheat varieties.

To classify wheat cultivars based on morphological features, artificial neural networks were used. ANN was trained using 394 samples and validated by 98 samples. Network evaluation was done using testing data (164 samples). Ten morphological features were used as input data. The best ANN was determined based on trial and error method. The number of hidden layers and the number of neurons in hidden layers are important factors in developing a MLP. Results showed that the developed MLP with one hidden layer in this study could successfully recognize most wheat cultivars. When the number of hidden layer was increased from one to two layers, the MLP's performance did not improve (Table 2), this can be due to over fitting problem in which the MLP memorize the input patterns. Therefore, ANN architecture with one hidden layer was selected for further investigation.

To find the optimum number of neurons in the hidden layers, trial and error method was used. Mean squared error (MSE) of test data was used for comparing MLP models with different number of neurons in hidden layer. The best model was selected based on the minimum value of MSE. The performance of different MLP models is presented in Table 2. Trends of network MSE with number of neurons in hidden layer shows a decreasing pattern up to 26 neurons and increasing pattern for MSE after 26 neurons. Then to prevent overtraining, the best number of neurons in the hidden layer is 26. Hence the best topology of ANN was obtained as 9-26-4.

Confusion matrix and classification accuracy of ANN classifier are presented in table 3 and table 4, respectively. The overall accuracy of wheat cultivar identification was 85.7%. The acceptable classification accuracies were obtained for 'Gohar', 'Dehdasht' and 'Seimareh' cultivars (Table. 4). 'Gohar' and 'Dehdasht' cultivars were classified with the highest accuracy and the lowest classification accuracy was obtained for 'Koohdasht' and 'Seimareh' cultivars. It may be because there were several morphological features that were able to make a significant difference between 'Gohar' and 'Dehdasht' with other cultivars (Table 5). For this reason, ANN was able to detect them with high accuracy. Whereas, a significant difference was not observed between 'Koohdasht' and 'Seimareh' cultivars for robust features. As an example, aspect ratio, compactness and circularity as the most robust features were not able to difference 'Koohdasht' and 'Seimareh' cultivars from each other (Table 5). Therefore, recognition of 'Koohdast' cultivar, especially, faced with difficulty and the lowest accuracy of classification was obtained for them.

Table 2. Performance of MLP for different number of neurons intended for validation data.

Number of hidden layers	Neurons in first hidden layer	Neurons in second hidden layer	MSE
1	2	0	0.071
1	4	0	0.061
1	6	0	0.059
1	8	0	0.048
1	10	0	0.053
1	12	0	0.053
1	14	0	0.062
1	16	0	0.055
1	18	0	0.058
1	20	0	0.043
1	22	0	0.053
1	24	0	0.055
1	26	0	0.041
1	28	0	0.057
1	30	0	0.058
2	2	2	0.064
2	4	4	0.058
2	6	6	0.056
2	8	8	0.056
2	10	10	0.050
2	12	12	0.055
2	14	14	0.048
2	16	16	0.054
2	18	18	0.060
2	20	20	0.055
2	22	22	0.052
2	24	24	0.054
2	26	26	0.061

Overall, the results show the developed computer vision system can identify wheat cultivars satisfactorily. Although the morphological features were not suitable for describing 'Koohdasht' cultivar. Further studies must be explored for finding suitable features such as, colour, texture, wavelet and Fourier transform for improving classification accuracy of 'Koohdasht' cultivar.

Table 3. Confusion matrix of ANN classifier for wheat cultivar identification.

Output/ Desired	Gohar	Dehdasht	Koohdasht	Seimareh
Gohar	42	0	0	0
Dehdasht	0	42	0	1
Koohdasht	0	0	25	4
Seimareh	0	2	16	32

Table 4. The ANN classification accuracy for wheat cultivar identification.

Wheat Cultivar	Accuracy
Gohar	100
Dehdasht	95.45
Koohdasht	60.97
Seimareh	86.48
Average	**85.72**

CONCLUSION

Combination of image processing and artificial neural network was successful in classification of 'Gohar', 'Dehdasht' and 'Seimareh' wheat cultivars. The highest classification accuracy was achieved for 'Gohar' cultivar. Morphological features could not describe significant difference between 'Koohdasht' and other cultivar well. The aspect ratio, compactness and circularity were the most important morphological features for wheat class identification.

Table 5. Mean comparison of morphological features for wheat cultivars.

Varieties	Perimeter	Major Axis	Minor Axis	Eccentricity	Solidity	Aspect Ratio	Area Ratio	Circularity	Compactness
Gohar	20.14 [b†]	7.50[d]	3.81[a]	0.85[c]	0.97[a]	1.97[c]	0.78[a]	18.16[c]	0.49[c]
Dehdasht	21.73 [a]	8.73[a]	3.29[b]	0.92[a]	0.96[a]	2.67[a]	0.78[a]	21.04[a]	0.6[b]
Koohdasht	21.36 [a]	7.74[c]	3.78[a]	0.87[bc]	0.96[a]	2.05[bc]	0.78[a]	19.9[b]	0.63[a]
Seimareh	22.12 [a]	8.17[b]	3.86[a]	0.88[b]	0.96[a]	2.11[b]	0.78[a]	19.8[b]	0.63[a]

† Treatments with the same letter don't have any significant difference at %5 probability level.

REFERENCES

Arefi, A., A.M. Motlagh and A. Khoshroo. 2011. Recognition of weed seed species by image processing. Journal of Food, Agriculture and Environment. 9: 379-383.

Cheng, H.D., X.H. Jiang, Y. Sun and J. Wang. 2001. Colour image segmentation: advances and prospects. Pattern Recognition. 34: 2259-2281.

Choudhary, R., J. Paliwl and D.S. Jayas. 2008. Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. Biosystem Engineering. 99: 330-337.

Dubey, B.P., S.G. Bhagwat, S.P. Shouche and J.K. Sainis. 2006. Potential of artificial neural networks in varietal identification using morphometry of wheat grains. Biosystems Engineering. 95: 61–67.

Khoshroo, A., A. Keyhani, S. Rafiee, R.A. Zoroofi and Z. Zamani. 2009. Pomegranate quality evaluation using machine vision. Proceedings of the First International Symposium on Pomegranate and Minor Mediterranean Fruits, Adana, Turkey. pp: 347-352.

FAO. 2011. http://faostat.fao.org/site/339/default.aspx.

Majumdar, S. and D.S. Jayas. 2000a. Classification of cereal grains using machine vision: I. Morphology models. Transaction of American Society of Agricultural Engineering. 43: 1669-1675.

Majumdar, S. and D.S. Jayas. 2000b. Classification of cereal grains using machine vision: II. Colour models. Transaction of American Society of Agricultural Engineering. 43: 1677-1680.

Majumdar, S. and D.S. Jayas. 2000c. Classification of cereal grains using machine vision: III. Texture models. Transaction of American Society of Agricultural Engineering. 43: 1681-1687.

Majumdar, S. and D.S. Jayas. 2000d. Classification of cereal grains using machine vision: IV. Combined morphology, colour, and texture models. 43: 1689-1694.

Mansing, J.S. 2010. Characterization of wheat (*Triticum* spp.) genotypes through morphological, chemical and molecular markers. A thesis submitted to the University of Agricultural Sciences, Dharwad. p: 1.

Masoumiasl, A., R. Amiri-Fahliani and A. Khoshroo. 2013. Some local and commercial rice (*Oryza sativa* L.) varieties comparison for aroma and other qualitative properties. International Journal of Agriculture and Crop Sciences. 5: 2184-2189.

Shouche, S.P., R. Rastogi, S.G. Bhagwat and K.S. Jayashree. 2001. Shape analysis of grains of Indian wheat varieties. Computers and Electronics in Agriculture. 33: 55-76.

Utku, H. 2000. Application of the feature selection method to discriminate digitized wheat varieties. Journal of Food Engineering. 46: 211-216.

Visen, N.S., D.S. Jayas, J. Paliwal and N.D.G. White. 2004. Comparison of two neural network architectures for classification of singulated cereal grains. Canadian Biosystems Engineering. 46:3.7-3.14.

Zapotoczny, P., M. Zielinska and M. Nitab. 2008. Application of image analysis for the varietal classification of barley: Morphological features. Journal of Cereal Science. 48:104-110.