DESIGN, DEVELOPMENT AND EVALUATION OF A GRADING SYSTEM FOR PEELED PISTACHIOS

Induja C[1], Kajal Shinde U[2], Shrinisha S S[3], Dr.Vinoth Chakkaravarthy G[4]

[1,2,3] Student(UG), VIII Semester B.E Dept. of Computer Science & Engineering, Velammal College of Engineering & Technology Madurai, Tamilnadu, India
[4] Assistant Professor, Dept. of Computer Science & Engineering, Velammal College of Engineering & Technology Madurai, Tamilnadu, India

---***---

Abstract - *In this study, an intelligent system based on combined machine vision (MV) and Support Vector Machine (SVM) was developed for sorting of peeled pistachio kernels and shells. A color CCD camera was used to capture images. The images were digitalized by a capture card and transferred to a personal computer for further analysis. Initially, images were converted from RGB color space to HSV color ones. For segmentation of the acquired images, H-component in the HSV color space and Otsu thresholding method were applied. A feature vector containing 30 color features was extracted from the captured images. A feature selection method based on sensitivity analysis was carried out to select superior features. The selected features were presented to SVM classifier. Various SVM models having a different kernel function were developed and tested. The SVM model having cubic polynomial kernel function and 38 support vectors achieved the best accuracy (99.17%).*

***Key Words*: Preprocessing,Feature Extraction, Segmentation, Classification, SVM,Ostu.**

1.INTRODUCTION

Pistachio nut is the top non- petroleum export of Iran that accounted up to 60% of global pistachio market.. The methods extend from manual-machine grading, where the features are determined manually, under laboratory conditions to machine vision systems for automated high-speed fruit sorting Among many available methods for quality evaluation of the crops, machine vision (MV) systems have proven to be the most powerful . A MV system consists of two main parts: hardware and software parts. These systems like the human eye are strongly influenced by lighting system. This part has a significant effect on the quality and resolution of the captured images and also considerably affects overall performance and efficiency of the MV system.

Computer-generated artificial classifiers that are intended to mimic human decision making for product quality have recently been studied, intensively . Omid et al designed and evaluated an intelligent sorting system for open and closed-shell pistachio nuts. The system included a feeder, an acoustical part, an electronic control unit, a pneumatic air-rejection mechanism and ANN classifier. The recognition was based on combined PCA of impact acoustics and ANN classifier. To generate useful features, both time and frequency-domain analysis of recorded sound signals were performed. In a recent study, a new method based on MV system was developed for egg volume prediction. A full-automatic system has been developed to remove the unwanted materials from pomegranate seeds and classify the seeds into four classes. Another research proposed an algorithm based on image processing for grading chestnuts. The algorithm was successfully applied on the online sorting systems. Mustafa et al. developed a sorting and grading system based on image processing and Support Vector Machine (SVM). The developed system captured fruit's image using a regular digital camera. Then, the image was transmitted to the processing level where feature extraction, classification and grading was done using MATLAB.

They concluded that the system was enough to use for classifying and grading the different varieties of rice grains based on their interior and exterior quality. Combined artificial neural network with machine vision to identify five classes of almond according to visual features. The images of five classes of almond including normal almond, broken almond, double almond, wrinkled almond and shell of almond were acquired, segmented by Otsu's thresholding method and classified by ANN. Olgun et al. developed wheat grain classification system by machine vision and SVM classifier.

According to performed studies on production and pro-cessing steps of pistachio kernels (PK), after cracking of close pistachios some cracked pistachio shells (PS) remain. The purpose of this study was to develop an automated system for sorting of PKs from unwanted PSs by using MV technique and SVM classifier.

2. MATERIALS AND METHODS

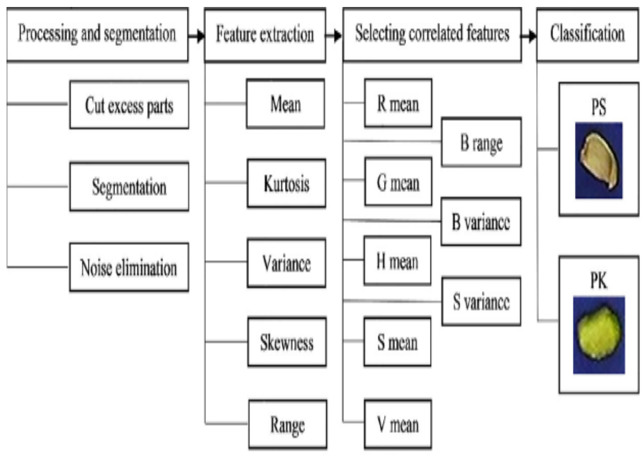

Fig 2a Framework of the system

2.1 Image Segmentation

Segmentation is the primary stage of image processing tasks. Segmentation strongly affects accuracy of image processing that so should be carefully implemented. Thresholding is an essential and important method for image segmentation . One of the widely used methods thresholding is Otsu's technique . In this research we also used Otsu method to separate the pistachio kernel from background. Primary tests indicated that the best color component for segmentation purpose was H-component of the images in HSV color space that successfully separated objects from background. However, if this component was directly used for the segmentation, some parts of the samples were also considered as background. So it was necessary to per-form histogram processing. Accordingly, as shown in , pixels values above 0.85 were replaced with zero. After that,the H-component was suitable for the segmentation. In Pre-liminary tests, we found that after segmentation, objects with less than 400 pixels in the segmented images were noises. Therefore, these objects were eliminated.

2.2 Feature Extraction

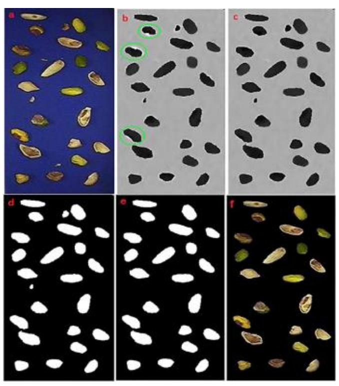

Fig 2b Steps involved in the sorting of pistachios

Color, texture, size and shape are the most important visual features of an object in the image. As shown in Fig 2b, there is a suitable contrast between the two classes, PK and PS, in terms of color. Logically, it is better to use only superior features for classification in the online systems to decrease processing time and speed up the system. The advantage of color space is clear distinction that is set in the relevant parameters of color between classes. This clear distinction can play a vital role in pistachio sorting, accurately. To verify the differences between PKs and PSs in different color spaces, various color descriptors were applied. So, color as one of the most important indicators was used to create the feature vector. Indicators usually are used to extract the color features are mean, variance, skew, kurtosis and range are evaluated using formulas. For this purpose, the five indices for each sample in the images were extracted from the components of RGB and HSV color spaces. Accordingly, a total of 30 features were obtained for each sample. To improve the performance and speed of the sorter, it was necessary to choose the features that were more effective in determining the correct output. The best feature selection would speed up the system decision making task which results an increased performance of the sorter. A good feature has the property that it is similar among the same class of objects and dissimilar among different classes of objects. Different feature selection techniques and classification methods have been developed and introduced by researchers who work in the field of artificial intelligence. Some of feature selection methods are Forward Feature Selection, Backward Feature Selection, PCA, Genetic Algorithm, etc.

2.4 Classification with support vector machines (SVMs)

The SVM is a supervised learning method that is widely used for classification. Although initially designed for binary classification, the basic SVM approach can be extended for the multi-class classification task.

A detailed mathematical explanation of SVM can be found in Vapnik's paper. Here, only some of the main features are presented. A SVM performs classification by map-ping input vectors into a higher-dimensional space and constructing a hyper-plane that optimally separates the data in the higher-dimensional space. Given a training set of instance-label pairs (x_i, y_i), in the classification type of SVM, the training involves the following optimization problem:

$$Min_{\omega,b,\xi}\frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}\xi_i$$

Subject to:

$$y_i(\omega^T\varphi(x_i)+b) \geqslant 1-\xi_i \text{ and } \xi_i \geqslant 0,\ i=1,\ldots,n$$

where $x_i = (x_1, x_2, \ldots, x_n)$, and $y_i = 1$ if x_i is in class 1, and y_i = _1 if x_i is in class 2. C > 0 is the penalty parameter of the error term., x is the vector of coefficients, b is a constant and n_i is a parameter for handling non-separable data (inputs). In the new space the data are considered as linearly sepa-rable.

There are several different kernel functions used in constructing SVM models. In this study, various kernel function such as linear, polynomial, and radial basis function (RBF) kernels were examined.

The following steps were implemented for selecting the most optimum SVM model:

1.Dividing data: 240 samples for training and validation and 120 samples for testing.

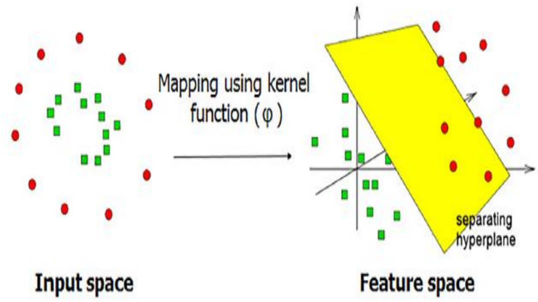

Fig 2c Principle of SVM

2.Designing and developing various SVM models with different kernel functions (linear, quadratic, polynomial, and radial basis).

3.Selecting the best model based on the highest accuracy (AC) at the testing stage and the minimum number of support vectors, since by selecting SVM with minimum number of support vectors, the processing speed would be increased and makes the algorithm more appropriate to use in the online system.

2.5 Performance analysis

To examine performance of the online system, two statistical indices including correct classification rate (CCR) and accuracy (AC) were used.

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$CCR = \frac{N_{Right}}{N_C}$$

where TP, TN, FP and FN are the numbers of true positives, true negatives, false positives and false negatives, respectively. For example, for PK class, TP is the number of samples in PK class when they actually are PK class, and TN is the number of samples in PS class when they are actually related to this class. Similarly, FP is the number of samples in PS class when they are actually related to PK class, and FN is the number of samples in PS class when they are actually related to PK class. In the CCR equation, N_{Right} and N_C refer to the number of samples correctly classified and total number of samples in that class, respectively.

3. RESULTS AND DISCUSSION

Fig. 3a shows results of SA on the 30 color features. 8 features that had the highest effect on identifying the output classes were selected for further analysis. The selected features were highlighted (bold) in Fig. 3a.

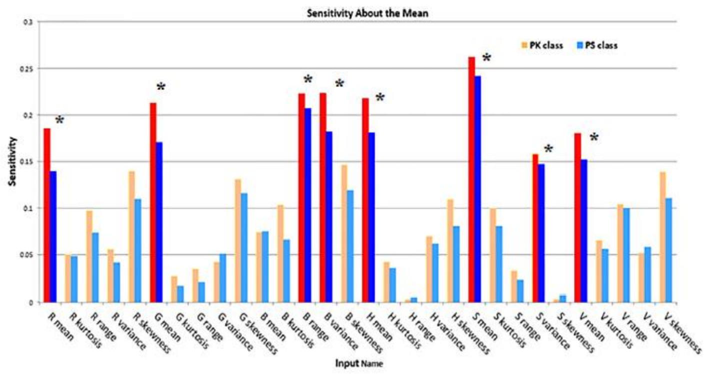

Fig 3a The eight superior features

Different models of SVM were developed, the best results for each model are presented in Table 2. The best kernel function was found to be the polynomial with order of 3 and 38 support vectors.

Table 2 – Results of pistachio classification by SVMs using different kernel functions.

Kernel function	SVs	AC (%)
Linear	47	95.83
Polynomial (degree = 2)	39	96.67
Polynomial (degree = 3)	38	99.17
Radial basis (Sigma = 0.4)	210	96.67

Accuracy (AC) are obtained from the confusion matrix during testing phase, and SVs are the number of support vectors.

The confusion matrix for cubic polynomial kernel function and the values for AC and CCR of classes are presented in Table 3. The excellent results were obtained for offline classification, i.e., the accuracy for training and validation data set and testing data set were 99.58% and 99.17%, respectively.

Table 3 – The confusion matrix obtained from the SVMs with the linear kernel function and performance evaluation by Accuracy and CCR.

Desired	Training and validation		Testing	
Predicted	PK	PS	PK	PS
PK	120	1	59	0
PS	0	119	1	60
CCR (%)	100	99.17	98.33	100
Accuracy (%)	99.58		99.17	

4. CONCLUSION

This work described an engineering solution for automatic sorting of the pistachio kernels from the unwanted shells. Superior features were extracted by means of SA method. Results of offline classification showed that SVM classifier with the cubic polynomial kernel function was accurate and efficient to be used for the online system. It was found that the capacity of the sorter was limited by the hardware parts of the system. This was mainly due to the distance between samples on the conveyor belt to reach 8 mm apart and pneumatic valves response time. The sorter has the capacity of 22.74 kg/h with the accuracy of 94.33%.

REFERENCES

1) Dashti G, Khodavardizadeh M, Rezaee MR. Analysis of pistachio's comparative advantages and global export market structure. J Dev Agric Econ 2010;24:99–106.

2) Mustafa NB, Ahmed SK, Ali Z, Yit WB, Abidin AA, Sharrif ZA. Agricultural produce sorting and grading using support vector machines and fuzzy logic. In: InSignal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on 2009 Nov 18 (pp. 391–396). IEEE.

3) Kumara B, Jinap S, Che Man YB, Yusoff MSA. Note: comparison of colour techniques to measure chocolate fat bloom. Food Sci Technol Int 2003;9:0295–299.

4) Shahabi M, Rafiee S, Mohtasebi SS, Hosseinpour S. Image analysis and green tea color change kinetics during thin-layer drying. Food Sci Technol Int 2014;20:465–76.

5) Jha NS. Nondestructive evaluation of food quality. Heidelberg, Dordrecht, London, New York: Springer; 2010.

6) Kumar-patel K, Kar A, Jha SN, Khan M. Machine vision system: a tool for quality inspection of food and agricultural products. J Food Sci Technol 2012;49:123–41.

7) Kavdir I, Guyer ED. Evaluation of different pattern recognition techniques for apple sorting. Biosyst Eng 2008;99:211–9.

8) Omid M, Mahmoudi A, Omid MH. Development of pistachio sorting system using principal component analysis (PCA) assisted artificial neural network (ANN) of impact acoustics. Expert Syst Appl 2010;37:7205–12.

9) Soltani M, Omid M, Alimardani R. Egg volume prediction using machine vision technique based on Pappus theorem and artificial neural network. J Food Sci Technol 2014;52:3065–71.

10) Mollazade K, Omid M, Arefi A. Comparing data mining classifiers for grading raisins based on visual features. Comput Electron Agric 2012;84:124–31.

11) Blasco J, Cubero S, Gomez-sanis J, Mira P, Molto E. Development of a machine for the automatic sorting of pomegranate. J Food Eng 2009;90:27–34.

12) Choudhary R, Paliwal J, Jayas DS. Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. Biosyst Eng 2008;99:330–7.

13) ElMasry G, Cubero S, Molto E, Blasco J. In-line sorting of irregular potatoes by using automated computer-based machine vision system. J Food Eng 2012;112:60–8.

14) Kaur H, Singh B. Classification and grading rice using multi-class SVM. Int J Sci Res Publ 2013;3:1–5.

15) Teimouri N, Omid M, Mollazadeh K, Rajabipour A. An artificial neural network-based method to identify five classes of almond according to visual features. J Food Process Eng 2015. http://dx.doi.org/10.1111/jfpe.12255. Olgun M Onarcan AO, Ö zkan K, Isik S, Sezer O, Ö zgisi K, et al. Wheat grain classification by using dense SIFT features with SVM classifier. Comput Electron Agric 2016;122:185–90.

16) MathWorks. Matlab User's Guide. The MathWorks Inc., USA; 2014.

17) Gonzalez RC, Woods RE, Eddins SL. Digital image processing using MATLAB. second ed. SI: Gatesmark Pub; 2009.

18) Otsu N. A threshold selection method from gray-level histograms. IEEE Tran Syst Man Cybern 1979;9:62–6.

19) Jackman P, Sun DW, Du CJ, Allen P, Downey G. Prediction of beef eating quality from colour, marbling and wavelet texture features. Meat Sci 2008;80:1273–81.

20) Rahman A, Murshed M. Feature weighting and retrieval methods for dynamic texture motion features. Int J Comput Intell Syst 2009;2(1):27–38.

21) Rahman A, Murshed M, Dooley LS. Feature weighting methods for abstract features applicable to motion based video indexing. In Information Technology: Coding and Computing, 2004. In: Proceedings. ITCC 2004. International Conference on 2004 Apr 5, vol. 1. IEEE, pp. 676–680.

22) Cristianini N, Shawe-Taylor J. An introduction to supporvector machines. 1th ed. Cambridge University Press; 2000.

23) Teimouri N, Omid M, Mollazadeh K, Rajabipour A. A novel artificial neural networks assisted segmentation algorithm for discriminating almond nut and shell from background and shadow. Compt Electron Agric 2014;1(05):34–43.

24) Vesali F, Omid M, Kaleita A, Mobli H. Development of an android app to estimate chlorophyll content of corn leaves based on contact imaging. Comput Electron Agric 2015;116:211–20.