Nasirahmadi A, Miraei Ashtiani S-H.

[Bag-of-Feature model for sweet and bitter almond classification.](#)

Biosystems Engineering 2017, 156, 51-60.

Copyright:

© 2017. This manuscript version is made available under the [CC-BY-NC-ND 4.0 license](#)

DOI link to article:

<https://doi.org/10.1016/j.biosystemseng.2017.01.008>

Date deposited:

04/05/2017

Embargo release date:

10 February 2018

This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

1 **Bag-of-Feature model for sweet and bitter almond classification**

2
3 **Abozar Nasirahmadi^{*a}, Seyed-Hassan Miraei Ashtiani^b**

4
5 ^a School of Agriculture, Food and Rural Development, Newcastle University, Newcastle upon Tyne NE1 7RU,
6 UK

7
8 ^b Department of Biosystems Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad,
9 Iran

10 11 **Abstract**

12 The use of computer vision techniques in post-harvest processing of agricultural products has
13 increased considerably in recent years due to their non-destructive and rapid monitoring
14 abilities. Image processing, combined with pattern recognition, has been applied in in fruit
15 sorting and classification. In this study, a Bag-of-Feature (BoF) model is used for the
16 classification of 20 sweet and bitter almond varieties. Harris, Harris-Laplace, Hessian,
17 Hessian-Laplace and Maximally Stable Extremal Regions (MSER) keypoint detectors along
18 with a Scale Invariant Feature Transform (SIFT) descriptor are used in the BoF model. The k-
19 means clustering method is applied for building a codebook from keypoint descriptors. The
20 performance of 3 classifiers, which were k-Nearest Neighbour (k-NN), linear and chi-square
21 Support Vector Machine (L-SVM and Chi-SVM, respectively) were compared using
22 classification results in the model. It was observed that the Chi-SVM classifier outperformed
23 the k-NN and L-SVM classifiers. Using the BoF model, it was possible to detect and classify
24 sweet and bitter varieties with high overall accuracy.

25

*Corresponding author: abozar.nasirahmadi@ncl.ac.uk, a.nasirahmadi@gmail.com

Keywords: Almond, Bag-of-Feature, Classification, Keypoint, Machine vision

1. Introduction

Almonds (*Prunus dulcis*) contain numerous bioactive and nutritional ingredients (e.g. vitamin E, fibre, arginine, polyunsaturated and monounsaturated fatty acids). In 2014, worldwide almond production was about 2.69 million tonnes and the five top producing countries were the USA, Australia, Spain, Iran and Morocco (FAOSTAT, 2014).

Based on taste, almonds can be divided into two major species: bitter and sweet almonds (Borràs, Amigo, van den Berg, Boqué, & Busto, 2014). Usually, sweet almonds are used in the food industries or for human consumption without any pre-processing (Loghavi, Sourì, & Khorsandi, 2011), whereas bitter almonds are toxic due to their cyanogenic glycosides and cannot be consumed without removal of toxic contents. However, the essential oil of bitter almonds is used in the pharmaceutical and cosmetic industries (Loghavi et al., 2011). Sometimes, bitter and sweet almonds are unintentionally mixed during post-harvest processing, which gives an undesirable experience to consumers. Furthermore, a bitter almond contains 4 to 9 mg of highly toxic hydrogen cyanide, and consumption in large amounts could lead to death (Volpi, 2016). It is therefore important to distinguish bitter and sweet almonds due to health risks of cyanide intake and likely economic losses from reduced value (Toomey, Nickum, & Flurer, 2012).

There are several techniques for recognition and analysis of cyanogenic glycosides, which can serve as criteria for separation of sweet and bitter almonds, including high-performance liquid chromatography (Dicenta et al., 2002), micellar capillary electrophoresis (Campa et al., 2000) and thin layer chromatography (Zhao, 2012). Although these methods are highly accurate in measurement, and therefore sweet/bitter almond recognition, they are complex, expensive, time-consuming and destructive (Borràs et al., 2014). Two expert systems, i.e.

human vision and machine vision, are used for recognition and classification of agricultural products. Human vision-based inspection and classification methods have received less attention due to serious weaknesses, e.g. high costs, low speed, need for experienced work force, low efficiency and accuracy, particularly for crops with highly similar cultivars like almonds. In contrast, machine vision has been an important technique for a wide variety of applications in agriculture, post-harvest processing and food engineering. This technique is an alternative, cheap and non-contact method that can replace human observation and other destructive methods of food assessment.

There are several studies in the literature where computer vision, along with pattern recognition (expert systems) methods, has been applied for recognition and classification of agricultural products in recent years. For instance, shape, colour and textural features have been used for wheat kernel detection (Shrestha, Kang, Yu, & Baik, 2016), rice and paddy classification (Chaugule & Mali, 2016), classification of beans and asparagus (Nasirahmadi & Behrooz-Khazaei, 2013; Donis-González & Guyer, 2016) and barley variety detection (Szczypiński, Klepaczko, & Zapotoczny, 2015). However, no studies have yet been carried out on the topic of detection of sweet and bitter almond varieties by means of image processing as a non-destructive method. To differentiate sweet and bitter almonds, Micklander, Brimer, and Engelsen (2002) and Borràs et al. (2014) used Raman and near-infrared spectroscopy, respectively. They used data on different molecular structures of samples to separate bitter and sweet almonds. Despite their strengths, these techniques have drawbacks. Fluorescent compounds usually interfere with Raman spectroscopy, resulting in no suitable Raman spectra. In addition, during the laser measurement process, the sample might be altered or destroyed (Thygesen, Løkke, Micklander, & Engelsen, 2003). The weaknesses of the near-infrared spectroscopy method are that it is time-consuming and needs

a calibration procedure, in addition to complex analysis and data interpretation (Giokas, Thanasoulas, & Vlessidis, 2011).

In addition to the nutritional values of almond and widespread use in various industries (i.e. food, medical and cosmetics) the economic importance of this product is another reason to choose this crop for study. The world trade export value of almonds has reached US\$ 4.7 billion by 2014 (FAOSTAT, 2014). In this study, an algorithm based on a Bag-of-Feature (BoF) model for sweet and bitter almond variety detection is proposed. Over the last years, in order to carry out object recognition and classification, researchers in the machine vision field introduced the BoF model (Yoo & Kim, 2013). The model has been used due to its low computational cost and its robustness (Tamaki et al., 2013). The BoF model is based on image keypoint detectors, an image descriptor, visual dictionary and classifiers, which are addressed in the following paragraphs. Recently, some keypoint detectors and descriptors have been applied in agricultural-based research for detection and classification of products/objects using different classifiers. Thanks to the capability of high dimensional datasets classification and ability to well address the over-fitting problem (Lei et al., 2017) Support Vector Machine (SVM) classifier has been used in most of the studies recently. Five orchard insect species were classified by means of local features, a scale invariant feature transform (SIFT) descriptor and six classifiers by (Wen, Guyer, & Li, 2009). Similarly, in order to automatic fruit recognition in images, a Bag-of-Words along with a SVM classifier model was proposed by (Song et al., 2014). In another study by applying dense SIFT features and SVM, 40 different wheat grain varieties were classified with a satisfactory accuracy rate (Olgun et al., 2016). Pires et al. (2016) introduced a high accuracy method based on image local descriptors and SVM classifier for detecting soybean disease.

Although these studies have concentrated on recognition and classification of products by image analysis, no specific study for the detection of almond varieties has been reported.

Addressing the challenges of categorising sweet and bitter almonds with no need for sample pre-processing is a key factor in introducing this model into design of automatic sorting systems. Hence, we attempt here to employ a BoF model followed by k-Nearest Neighbour (k-NN) and SVM classifiers, for classification of 20 bitter and sweet almonds. The remaining sections are organised to give a description of the different steps in the BoF model, the almond varieties used in this work, the results and discussion of the experiment.

2. Material and methods

2.1. Image acquisition

Twenty varieties comprising, 8 bitter almonds: Ahmadabad, Talkhe-Kaghazi, Mehrizi, ME-34, Koohi, HC-9, Dehshir, Bahadoran and 12 sweet almonds: Sangi-12, Najafabadi, Sangi-05, MT-101, Shahrood-13, Harir, PA-C-9, Arzhan-Y, Mostofi, Dirgol-A-2, Badamak, Sefid were randomly selected from an almond orchard in a commercial farm from Yazd, Iran (Fig. 1). A fluorescent lamp with 36 W and of around 500 mm diameter was mounted above the samples to provide light during the capturing and colour (RGB) images were acquired using a commercial CCD camera (Nikon D3200). The camera was fixed at 300 mm above the samples with its lens pointing downward to acquire top view images with an original resolution of 4320×3240 pixels which were then cropped to 1024×1024 pixels resolution. It is crucial to detect background and foreground of an image because the background usually contains disturbing visual information that can affect the performance of models (Feng, Bhanu, & Heraty, 2016). Therefore, multi-threshold methods, as described by (Xiao-bo, Jie-wen, Yanxiao, & Holmes, 2010) were applied for background removal using the Image Processing Toolbox of MATLAB[®] (the Mathworks Inc., Natick, MA, USA).

2.2. Bag-of-features model

In order to detect the appearance of the different almond classes, a BoF model was applied. The BoF framework represents an image as a histogram of representative local features extracted from the image and called visual words or codewords (Yoo & Kim, 2013), and a set of codewords is generally called a visual vocabulary or a codebook (Tamaki et al., 2013). Figure 2 represents an overview of the recognition process with BoF. Firstly, keypoints are extracted from the training images and each keypoint is coded in an image descriptor. Then, these are clustered with vector quantisation to represent an image with a histogram of visual codewords. As all the features are mapped to the codebook, an image can be represented by the frequency histogram of BoF of the visual words in the codebook. Finally, classifiers are used to perform image classification. In all, BoF can be grouped into three components: detection and description of keypoints, codeword representation, and classification (Tamaki et al., 2013).

2.2.1. Detection and description of keypoints

In computer vision, extracting keypoints from local features (i.e. corner, edge, entropy, region, and curvature) helps to detect stable and informative points of interest. The keypoint detectors can be divided into three main groups: i) corner detectors such as Harris, Harris-Affine, Harris-Laplace and Smallest Univalued Segment Assimilating Nucleus (SUSAN), ii) region detectors, i.e. Maximally Stable Extremal Regions (MSER) and Intensity Based Regions (IBR), iii) blob detectors such as Hessian, Hessian-Affine, Fast-Hessian, Hessian-Laplace, Difference of Gaussian (DoG), Laplacian of Gaussian (LoG) (Tuytelaars & Mikolajczyk, 2008). In this study, at least one keypoint detector from each group was applied to select the best detector in almond images and these are described in the following paragraphs.

Harris detector: Harris and Stephens proposed a corner detector, named as the Harris detector, which is based on detecting changes in image intensity around a point using the second moment or autocorrelation matrix, where the matrix is composed of first order image derivatives (Tuytelaars & Mikolajczyk, 2008).

Harris-Laplace detector: A strong keypoint extraction method and the extracted features are invariant to image changes like rotation, scaling and translation. The Harris-Laplace detector first finds a set of images represented at different resolutions for reliable Harris detection, then for detecting keypoints an automatic scale selection procedure is applied to each corner by selecting the scale with a maximal LoG response (Wang, Niu, Yang, & Chen, 2012).

MSER detector: The MSER is one of the most stable methods of region detection. This detector is based on the principal to respond to extreme regions through a large range of thresholds that are stable and have luminance values that are considerably higher or lower than others in their surroundings (Gao et al., 2016).

Hessian detector: This detector captures important local properties of an image. It is based on the second derivatives of a Hessian matrix by searches for points where the determinant of the Hessian becomes maximal (Tuytelaars & Mikolajczyk, 2008).

Hessian-Laplace detector: This detector uses the Hessian matrix to locate points in space and the Laplacian function to compute their scale, and chooses interest points of the image where the determinants of the Hessian are maximal (Tuytelaars & Mikolajczyk, 2008).

For the detectors and descriptor, we used the Computer Vision System Toolbox[™] of MATLAB and the open source VLFeat library (Vedaldi & Fulkerson, 2008) in this study. Figure 3 shows the results of keypoint detection on two almond varieties using the five mentioned keypoint detectors in this study. The goals of keypoint descriptors are to reduce the computational complexity, extract some distinctive information around each point and

improve the robustness to image deformation (Wen et al., 2009). In this study, a well-known gradient-based feature descriptor which provides a robust keypoint descriptor is SIFT, proposed by Lowe (2004) for generating a set of local descriptors from an image. The SIFT descriptor is very stable to rotation, scale and affine changes (Lowe, 2004) and is one of the most widely used descriptors. In the SIFT descriptor, the orientations and gradient magnitudes are computed at each point of the local neighbourhood and weighted by a Gaussian function. The computation involves the size of 16×16 neighbourhood of each keypoint. Then it divides this into 4×4 sub-regions, so that an orientation histogram which represents eight cardinal directions is created for each sub-region based on gradient magnitude (Fig. 4). Finally, the keypoint descriptor is created that consists of 128 ($4 \times 4 \times 8$) elements.

2.2.2. Codeword representation

The next step is to build a codebook from the keypoint descriptor generated by the previous step. Since that there are thousands of keypoints that can be computed for each almond category, it becomes difficult to discover a set of common keypoints to one particular type of action. So, these large quantities of features are usually clustered into several hundreds of clusters using the k-means algorithm. The k-means clustering is the most popular and simplest partitioning algorithm which mainly uses the Euclidean distance for computing the distance between points and cluster centres to construct a histogram of visual words to represent the image (Jain, 2010). The number, or size, of the codebook has to be decided before execution of the clustering algorithm for better performance. The best codebook size depends on the size of the dataset, and in this study, the number of visual words was examined by accuracy for values ranging from 50 to 1100 for the almond dataset. The quantised vector was used to generate a histogram of words in the codebook. Each element of

the histogram shows the number of times that particular codewords has appeared in the image. The length of the vector is equal to the size of the codebook and classifiers then can be trained on image histogram vectors. Figure 5 illustrates some codeword histograms of four almond varieties (two bitter and two sweet). It can be seen that there are many differences between the histograms of varieties, even in the same group.

2.2.3. Classification

In this study two sets of classifiers have been tested: the distance based classifier (k-NN) and the learning based classifier (SVM).

The k-NN algorithm is one of the simplest but efficient machine learning algorithms and it is based on closest training examples in the feature space. The training process for this algorithm only consists of storing feature vectors and labels of the training images (Montoliu, Martín-Félez, Torres-Sospedra, & Martínez-Usó, 2015). K-NN scores the features of a category based on the category of k neighbours and it is commonly applied to evaluate and compare the performance of other classifiers due to its simplicity and classification effectiveness (Wang, An, Chen, Li, & Alterovitz, 2015). In k-NN classifier, generally various distance metrics apply to measure the distance between two samples, i.e. Manhattan, chi-square, Euclidean and Minkowski distances (Wang et al., 2015). In this study, 1NN (k=1) with the most common distance function for k-NN, which is a Euclidean-based distance, has been used (Montoliu et al., 2015; Wang et al., 2015).

SVM basically constructs a hyperplane in a high dimensional space that separates data into two categories for classification (Burges, 1998). The lower generalisation error of SVM, or a good separation, is achieved by a hyperplane that has the largest distance to the nearest training data point of any class (Montoliu et al., 2015). The test points are then mapped into that same space and predicted to belong to a class. In this work, the LibSVM library was used

(Chang & Lin, 2011) for SVM classifiers with linear (L-SVM) and chi-square kernel (Chi-SVM) functions. Cross-validation partitions the image dataset into complementary folds for ensuring each fold has the same proportion of each class (Montoliu et al., 2015; Pires et al., 2016). The SVM classifier needs to adjust some internal parameters, so a 10-fold cross-validation with 10 repetitions was used to evaluate the entire dataset in this study (Wen et al., 2009; Olgun et al., 2016). There were 20 almond varieties with 100 sample images each, giving a total of 2,000 images, and the images were randomly divided into a training set and a testing set. The number of the training and testing images were 1,700 and 300, respectively.

3. Results and discussion

In order to detect sweet and bitter varieties by a BoF model, five key feature detectors and a SIFT descriptor have been applied. Figure 6 demonstrates the average accuracy (10 repetitions) of five keypoint detectors, along with the standard deviation, in different codebook sizes. Comparing the accuracies of five keypoint detectors: Harris-Laplace, Hessian-Laplace, Hessian, Harris and MSER, for the k-NN classifier gave, 79.0% (± 2.1), 74.5% (± 1.4), 65.5% (± 2.5), 65.0% (± 1.6) and 57.2% (± 2.4) for 500, 500, 600, 600 and 700 visual words, respectively. For the L-SVM classifier the best accuracy obtained for 500 visual words was 86.7% (± 1.6) in Harris-Laplace, 80.0% (± 1.3) in Hessian-Laplace and 67.4% (± 1.6) in Hessian, while Harris and MSER gave 66.0% (± 1.5) and 61.2% (± 2.1) for 600 and 800 visual words, respectively. We can also observe that Harris-Laplace, Hessian-Laplace, Hessian and Harris require 500 visual words to achieve better accuracy in Chi-SVM, while MSER needs more (600). As shown in the figure, the Harris-Laplace detector showed a considerably higher value of accuracy than other detectors. According to Figs. 3 and 6, the Harris-Laplace detector covers pores on the almond shell in the whole varieties and more keypoints were extracted than by other detectors. Based on (Tuytelaars & Mikolajczyk, 2008)

the Harris-Laplace detector is a corner detector with partial blob detection ability and in recent years this detector has attracted great attention in object recognition, image retrieval and image registration due to high repeatability and localisation accuracy in comparison with other invariant-scale detectors (Alvarez-Betancourt & Garcia-Silvente, 2016).

In order to find the best codebook size in the k-means cluster, Table 1 was presented for the detector. It can be inferred from the table that the highest accuracy was obtained in a 500 codebook size for the detector. It can be seen that, when the size of codebook reaches to 500, the classification accuracy result is the best. The reason is that when the value of visual words is too small, the generated visual words are not enough to represent the entire classes of features and this causes deviation of performance. However, when the size of visual words is too large, some similar features can be divided into different classes. The dispersion of visual words leads to lack of generality, so it cannot describe the key features of those classes well, and leads to a decline in the classification performance (Zhao, Li, & Cang, 2015). As a result, the Harris-Laplace keypoint detector and codebook size of 500 were used for the BoF model of variety detection in this study.

By comparing the accuracy of classifiers in Table 1, it can be observed that SVM classifiers appear to perform better than the nearest neighbour classifier. Similar results were obtained by (Kazmi, Garcia-Ruiz, Nielsen, Rasmussen, & Andersen, 2015) for weed detection and (Wen et al., 2009) for orchard insect classification.

Two SVM kernel types show different performance, the chi-square kernel performs better than the linear one with test overall accuracy of 91 and 86.7%, respectively. Tamaki et al. (2013) observed the same result, because the Chi-SVM enjoys exponential computation and leads to higher discrimination performance than the linear kernel type (Huang et al., 2015).

Figure 7 shows the confusion matrices of the twenty classes, based on a run of the experiment. In the matrices, each row represents the varieties in an actual class, and each

column represents the varieties in a predicted class. The colour of each entry indicates the percentage of the classification and misclassification in the classifiers. A darker red series of cells diagonally across the matrices reflects accurate classification, as these cells represent the proportion of correctly predicted almond varieties. Furthermore, cells outside of the main diagonal describe the proportion of frames misclassified. It is obvious from the matrices that the number of misclassification with the k-NN classifier is higher than with the SVM classifier.

As can be seen from Fig. 7, with the Chi-SVM classifier eight out of twenty varieties, which are Koohi, HC-9, Sangi-12, Sangi-05, Shahrood-13, Arzhan-Y, Dirgol-A-2 and Badamak, achieved 100% correct classification. For these eight classes, no misidentification happened. The first 8 varieties (V1-V8) in the figure, belong to bitter almonds while other 12 (V9-V20) are sweet varieties. Between bitter samples, Dehshir has the lowest value of correct classifications and this sample was misidentified to be Bahadoran, Najafabadi and Shahrood-13 samples because of the similarity between these varieties. Similarly, among in sweet varieties, PA-C-9 showed the highest misclassification rate and was classified as two other samples (MT-101 and Talkhe-Kaghazi) which are highly similar to each other. Based on (Ghiassi and Burnley, 2010) classification error is the proportion of the number of samples misclassified to the total number of samples. For the purpose of finding the percentage of misclassified bitter into sweet almonds in Chi-SVM, classification error for each group was calculated. The results showed that around 2.5% of bitter samples were classified as sweet and 4.5% of sweet almond samples were misclassified into bitter categories. The results obtained from the performance of the Chi-SVM classifier show a good almond variety detection rate, with satisfactory overall accuracy (91%). According to the criteria of Fig. 7, some almond images were not recognised correctly and there were some false classifications. These errors sometimes occurred because the almonds were selected from a commercial farm

where no previous grading was carried out for the samples and some varieties might be so similar as to be 100% classified. It is clear that classification of these 20 almond varieties needs different methods from those of some other types of agricultural products due to similarities in the nature of their shape and colour. To date, no previous studies have been carried out on the classification of bitter and sweet almond varieties by image analysis. The technique proposed here can automatically classify varieties, even in samples with many similarities. The method could be a valuable tool to aid post-harvest processing and help the food packing industry to increase customer satisfaction, and reduce economic losses.

4. Conclusions

The study aimed to determine whether a BoF model could be effective for both sweet and bitter almond classification. The BoF model used in the study consisted of keypoint detectors and a SIFT descriptor, followed by k-NN, L-SVM and Chi-SVM classifiers. Our study compared the use of five keypoint detectors using an almond image data set composed of 2000 images. Among the keypoint detectors, the Harris-Laplace with a codebook size of 500 showed the highest accuracy among different classifiers. It was found that, Chi-SVM gave fewer misclassification among samples, however about 2.5 and 4.5% classification error was obtained between sweet and bitter almonds using this classifier.

The performance of the algorithm showed a high level of accuracy, so this method could contribute in the future as an important and economically feasible technique in almond post-harvest processing. In practice, the user can take photos from almond samples and apply the proposed approach. However, this method needs more study for enhancing classification accuracy. The combination of general features (i.e. colour, shape and texture) and some mentioned local features could be an alternative for future research into a fully automated machine vision technique in order to have the best classification performance.

Acknowledgment

The authors wish to thank the Yazd Agricultural Management Department for providing almond samples.

References

- Alvarez-Betancourt, Y., & Garcia-Silvente, M. (2016). A keypoints-based feature extraction method for iris recognition under variable image quality conditions. *Knowledge-Based Systems*, 92, 169–182.
- Borràs, E., Amigo, J. M., van den Berg, F., Boqué, R., & Busto, O. (2014). Fast and robust discrimination of almonds (*Prunus amygdalus*) with respect to their bitterness by using near infrared and partial least squares-discrimination analysis. *Food Chemistry*, 153, 15–19.
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121–167.
- Campa, C., Schmitt-Kopplin, Ph., Cataldi, T. R. I., Bufo, S. A., Freitag, D., & Kettrup, A. (2000). Analysis of cyanogenic glycosides by micellar capillary electrophoresis. *Journal of Chromatography B*, 739, 95–100.
- Chang, C.-C., & Lin, C.-J. (2011). Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(27), 1–27.
- Chaugule, A. A., & Mali, S. N. (2016). Identification of paddy varieties based on novel seed angle features. *Computers and Electronics in Agriculture*, 123, 415–422.
- Dicenta, F., Martínez-Gómez, P., Grané, N., Martín, M. L., León, A., Cánovas, J. A., & Berenguer, V. (2002). Relationship between cyanogenic compounds in kernels, leaves, and roots of sweet and bitter kernelled almonds. *Journal of Agricultural and Food Chemistry*, 50, 2149–2152.
- Donis-González, I. R., & Guyer, D. E. (2016). Classification of processing asparagus sections using color images. *Computers and Electronics in Agriculture*, 127, 236–241.
- FAOSTAT. (2014). *Food and agricultural commodities production*. Food and Agricultural Organization of the United Nations. <http://faostat3.fao.org>. Accessed on January 2017.

- Feng, L., Bhanu, B., & Heraty, J. (2016). A software system for automated identification and retrieval of moth images based on wing attributes. *Pattern Recognition*, 51, 225–241.
- Gao, Y., Shan, X., Hu, Z., Wang, D., Li, Y., & Tian, X. (2016). Extended compressed tracking via random projection based on MSERs and online LS-SVM learning. *Pattern Recognition*, 59, 245–254.
- Ghiassi, M., & Burnley, C. (2010). Measuring effectiveness of a dynamic artificial neural network algorithm for classification problems. *Expert Systems with Applications* 37, (4), 3118-3128.
- Giokas, D. L., Thanasoulas, N. C., & Vlessidis, A. G. (2011). Multivariate chemometric discrimination of cigarette tobacco blends based on the UV-Vis spectrum of their hydrophilic extracts. *Journal of Hazardous Materials*, 185, 86–92.
- Huang, S., Qi, L., Ma, X., Xue, K., Wang, W., & Zhu, X. (2015). Hyperspectral image analysis based on BoSW model for rice panicle blast grading. *Computers and Electronics in Agriculture*, 118, 167–178.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31, 651–666.
- Kazmi, W., Garcia-Ruiz, F., Nielsen, J., Rasmussen, J., & Andersen, H. J. (2015). Exploiting affine invariant regions and leaf edge shapes for weed detection. *Computers and Electronics in Agriculture*, 118, 290–299.
- Lei, B., Li, W., Yao, Y., Jiang, X., Tan, E-L., Qin, J., Chen, S., Ni, D., & Wang, T. (2017). Multi-modal and multi-layout discriminative learning for placental maturity staging. *Pattern Recognition*, 63, 719-730.
- Loghavi, M., Souri, S., Khorsandi, F. (2011). Physical and mechanical properties of almond (*Prunus dulcis* l. Cv. 7Shahrood). In *Paper presented at the 2011 ASABE annual Meeting Louisville, KY*.
<http://elibrary.asabe.org/abstract.asp?aid=37425>.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of Computer Vision*, 60, 91–110.

377 Micklander, E., Brimer, L., & Engelsen, S. B. (2002). Noninvasive assay for cyanogenic constituents
378 in plants by raman spectroscopy: Content and distribution of amygdalin in bitter almond (*Prunus*
379 *amygdalus*). *Applied Spectroscopy*, 56, 1139–1146.

380 Montoliu, R., Martín-Félez, R., Torres-Sospedra, J., & Martínez-Usó, A. (2015). Team activity
381 recognition in Association Football using a Bag-of-Words-based method. *Human Movement*
382 *Science*, 41, 165–178.

383 Nasirahmadi, A., & Behroozi-Khazaei, N. (2013). Identification of bean varieties according to color
384 features using artificial neural network. *Spanish Journal of Agricultural Research*, 11, 670–677.

385 Olgun, M., Onarcın, A. O., Özkan, K., Işık, Ş., Sezer, O., Özgişi K., Ayter, N. G., Başçiftçi, Z. B.,
386 Ardiç, M., & Koyuncu, O. (2016). Wheat grain classification by using dense SIFT features with
387 SVM classifier, *Computers and Electronics in Agriculture*, 122, 185–190.

388 Pires, R. D. L., Gonçalves, D. N., Oruê, J. P. M., Kanashiro, W. E. S., Rodrigues J. F., Machado, B.
389 B., & Gonçalves, W. N. (2016). Local descriptors for soybean disease recognition. *Computers and*
390 *Electronics in Agriculture*, 125, 48–55.

391 Shrestha, B. L., Kang, Y. M., Yu, D., & Baik, O. D. (2016). A two-camera machine vision approach
392 to separating and identifying laboratory sprouted wheat kernels. *Biosystems Engineering*, 147,
393 265–273.

394 Song, Y., Glasbey, C. A., Horgan, G. W., Polder, G., Dieleman, J. A., & van der Heijden, G. W. A.
395 M. (2014). Automatic fruit recognition and counting from multiple images. *Biosystems*
396 *Engineering*, 118, 203–215.

397 Szczypiński, P. M., Klepaczko, A., & Zapotoczny, P. (2015). Identifying barley varieties by computer
398 vision. *Computers and Electronics in Agriculture*, 110, 1–8.

399 Tamaki, T., Yoshimuta, J., Kawakami, M., Raytchev, B., Kaneda, K., Yoshida, S., Takemura, Y.,
400 Onji, K., Miyaki, R., & Tanaka, S. (2013). Computer-aided colorectal tumor classification in NBI
401 endoscopy using local features. *Medical Image Analysis*, 17, 78–100.

402 Thygesen, L. G., Løkke, M. M., Micklander, E., & Engelsen, S. B. (2003). Vibrational
403 microspectroscopy of food. Raman vs. FT-IR. *Trends in Food Science and Technology*, 14, 50–57.

- Toomey, V. M., Nickum, E. A., & Flurer, C. L. (2012). Cyanide and Amygdalin as indicators of the presence of bitter almonds in imported raw almonds. *Journal of Forensic Sciences*, 57, 1313–1317.
- Tuytelaars, T., & Mikolajczyk, K. (2008). Local invariant feature detectors: A survey. *Foundations and trends® in Computer Graphics and Vision*, 3, 177–280.
- Vedaldi, A., & Fulkerson, B. (2008). VLFeat: an open and portable library of computer vision algorithms. <<http://www.vlfeat.org/>>.
- Volpi, G. (2016). Demonstrating the presence of cyanide in bitter seeds while helping students visualize metal-cyanide reduction and formation in a copper complex reaction. *Journal of Chemical Education*, 93, 891–897.
- Wang, A., An, N., Chen, G., Li, L., & Alterovitz, G. (2015). Accelerating wrapper-based feature selection with K-nearest-neighbor. *Knowledge-Based Systems*, 83, 81-91.
- Wang, X. Y., Niu, P. P., Yang, H. Y., & Chen, L. L. (2012). Affine invariant image watermarking using intensity probability density-based Harris Laplace detector. *Journal of Visual Communication and Image Representation*, 23, 892–907.
- Wen, C., Guyer, D. E., & Li, W. (2009). Local feature-based identification and classification for orchard insects. *Biosystems Engineering*, 104, 299–307.
- Xiao-bo, Z., Jie-wen, Z., Yanxiao, L., & Holmes, M. (2010). In-line detection of apple defects using three color cameras system. *Computers and Electronics in Agriculture*, 70, 129–134.
- Yoo, S., & Kim, C. (2013). Background subtraction using hybrid feature coding in the bag-of-features framework. *Pattern Recognition Letters*, 34, 2086–2093.
- Zhao, C., Li, X., & Cang, Y. (2015). Bisecting k-means clustering based face recognition using block-based bag of words model. *Optik - International Journal for Light and Electron Optics*, 126, 1761–1766.
- Zhao, Y. Y. (2012). Amygdalin content in four stone fruit species at different development stages. *ScienceAsia*, 38, 218–222.

432

433 Table 1. Accuracy of Harris -Laplace detector for each classifier in different codebook sizes along
 434 with standard deviations (in parentheses).

435

Codebook size	k-NN	L-SVM	Chi-SVM
100	72.3% (± 1.3)	79.0% (± 2.3)	83.5% (± 2.0)
200	73.2% (± 2.1)	79.5% (± 2.4)	85.0% (± 2.1)
300	76.5% (± 2.2)	83.4% (± 1.2)	88.5% (± 1.4)
400	77.0% (± 2.8)	84.5% (± 1.8)	89.0% (± 1.7)
500	79.0% (± 2.1)	86.7% (± 2.0)	91.0% (± 1.6)
600	78.5% (± 1.5)	86.0% (± 1.5)	89.3% (± 2.5)
700	76.0% (± 1.8)	83.5% (± 1.6)	87.3% (± 2.6)
800	75.0% (± 2.4)	83.0% (± 2.1)	88.0% (± 2.4)
900	70.5% (± 1.6)	84.2% (± 2.4)	86.5% (± 2.2)
1000	68.3% (± 2.2)	82.1% (± 2.5)	85.2% (± 1.1)
1100	67.0% (± 2.3)	81.0% (± 2.2)	85.0% (± 2.4)

Figure Captions

Fig. 1. Almond varieties used in this study after background removing.

Fig. 2. Architecture of the BoF model.

Fig. 3. Example of keypoints detection using different detectors on two almond varieties, (top) bitter, (bottom) sweet.

Fig. 4. Example of graphical representation of the SIFT descriptor for an almond.

Fig. 5. Example of codeword histograms, (left) bitter and (right) sweet.

Fig. 6. Accuracy obtained from each classifier by five keypoint detectors in different codebook sizes.

Fig. 7. Confusion matrices of twenty variety classifications of three classifiers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

463 Fig .1.
464

Fig .2.

473 Fig .3.
474

Fig .4.

489 Fig .5.
490

491 Fig.6.

492

493

494

495

Fig .7.