DETECTION OF DRIED FIGS WITH BLACK MOLD BY USING HYPERSPECTRAL IMAGES

G. Ortaç, K. Taşdemir *

Antalya International University
Department of Computer Engineering
Universite Cd. 2, Dosemealti, Antalya

A. S. Bilgi, E. Durmuş, H. Kalkan

Suleyman Demirel University
Department of Computer Engineering
Cunur, Isparta, Turkey

ABSTRACT

Hyperspectral imaging systems have been recently popular for food quality and safety assessment, due to their ability to reflect unique spectral properties of materials at narrow band intervals. They help detection of contamination in foods such as dash, mold, crush, fungi. We propose such a system for effective detection of black molds in dried figs to avoid the high cost of manual process. The proposed system depends on finding the best discriminative spectral band and the optimum local spatial characteristics using forward feature selection with commonly used classifiers. The preliminary accuracies of upto 93.58% are promising for an operational system.

Index Terms— Food quality, forward feature selection, classification, dried figs, black mold, Aspergillus niger.

1. INTRODUCTION

In modern food industry, consumers demand food products with high quality and assurance of food safety. Thus, food quality and safety are major concerns for this industry [1], coupled with the requirements of low production costs. Meeting these challenges efficiently has become critical for food companies and suppliers [2].

A recently popular technique for food quality and safety assessment is the use of hyperspectral imaging analysis (HSIa) which depends on distinctive reflectance properties of the materials measured at many narrow bands and their spatial characteristics [3, 4]. HSIa has been used in several different applications. [5] use hyperspectral imaging for fast detection of POD (peroxidase) activity in tomato leaves with reflectance data in the visible and NIR spectral range and report acceptable prediction results with genetic algorithm-partial least squares (GA-PLS) at selected optimal wavelengths. [6] acquire reflectance images between 400-1000 nm to detect various common defects on orange surfaces by principal component analysis. [7] provide a reflectance technique to find the optical properties of fresh fruits and vegetables using diffuse reflectance technique with region between 500-1000 nm. [8] aim to detect bruises on apples by hyperspectral imaging between 900 and 1700 nm with principal component transform and minimum noise fraction transform. They show that the spectral region would be convenient for detecting bruises. Another study detects the bacterial contamination of meat and vegetables in the reflectance mode range 900-1700 nm by applying partial least squares to the logarithmic values [9]. Beside these studies, some studies use both reflectance and transmittance images for quality assessment. One of such research evaluates the internal quality of blueberries in reflectance and transmittance mode using partial least squares for soluble solids contents and firmness index [10] to conclude that the reflectance mode performs better than the transmittance mode. The study of the detection of fruit flies in pickled cucumber in [11] also uses integrated modes to achieve classification accuracies between 82% and 88% with reflectance mode by using partial least squares discriminant analysis.

In addition to the above mentioned fresh food sectors, another important type with increasing quality demands is the dried fruits market where many criteria should be investigated to determine the existence of hazardous materials (such as dash, mold, crush, fungi). Even though some kinds of these criteria can be detected by manual investigation, some kinds of these, especially those occurring inside of fruits (such as black molds in dried figs), cannot be entirely recognized by manual investigation. Although there are 23 different Aspergillus spp causing black molds, the common one is Aspergillus niger (existing in more than 0.2% of the figs). Even though Aspergillus niger exists inside the dried figs, it is observed that it may also damage outer-surface of figs and cause variation of fig shell (Fig. 1). Many studies have investigated decay of figs caused by A. niger [12].

Traditionally, the figs with black mold are manually selected by workers with a method called nailing where figs are controlled with a knurled skewer. This skewer is inserted from the bottom of fig and taken out to check whether any infection is remained on the skewer. Even though this method is useful, it has some disadvantages: black molds can be transferred from one fig to another by the skewer even if it is

*This work is funded in part by TUBITAK (The Scientific and Technological Research Council of Turkey) Grant No 113O878 and 113O879. We thank TARIŞ for providing the labelled figs for our experiments.

Fig. 1. Dried fig samples with damaged outer-surface because of black mold inside them.

cleaned at every step, and whole black molds cannot be detected by this method. In addition, this manual process requires an enormously high workforce and heavy processing time, resulting in high cost.

This study develops a hyperspectral imaging system operated in reflectance mode to automatically detect black mold on dried figs. The objective is to separate contaminated fig samples using variations of fig shell which occur because of Aspergillus Niger infection inside the fig. Based on a hyperspectral camera with a spectral region from 400 to 1000 nm (0.78nm intervals), we first determine the optimal wavelengths for infection detection. Secondly, we segment the fig image to work on local regions and to detect certain places of infections. Third we calculate the average intensity values of these local regions to extract features. We then select the optimum features based on forward feature selection and commonly used classifiers (linear discriminant, quadratic discriminant, support vector machine, logistic linear, and k-nearest neighbor), to achieve high accuracies. The preliminary results produce a promising accuracy of upto 93.58%. The rest of the paper is organized as follows. Section 2 proposes the system, Section 3 presents the experimental results and Section 4 concludes the paper.

2. PROPOSED METHOD

The proposed hyperspectral imaging system, outlined in Fig. 2 has five steps: image acquisition, image pre-processing, band selection with Fisher linear discriminant distance, image segmentation, feature selection and classification.

In the first step, image acquisition is performed by a system operated in a reflectance mode (400-1000 nm) with a 12-bit CMOS camera (model Specim Spectral Camera PFD-C-XX-V10E), a round conveyor belt and two 150 W tungsten halogen lamps connected to a dual fiber optic line positioned above the belt (Fig 3). The camera is set to run for 10 ms exposure time with the conveyor belt speed of 0.6 rpm and the samples are scanned line by line. Each sample is represented by a total of 784 intensity values for each wavelength

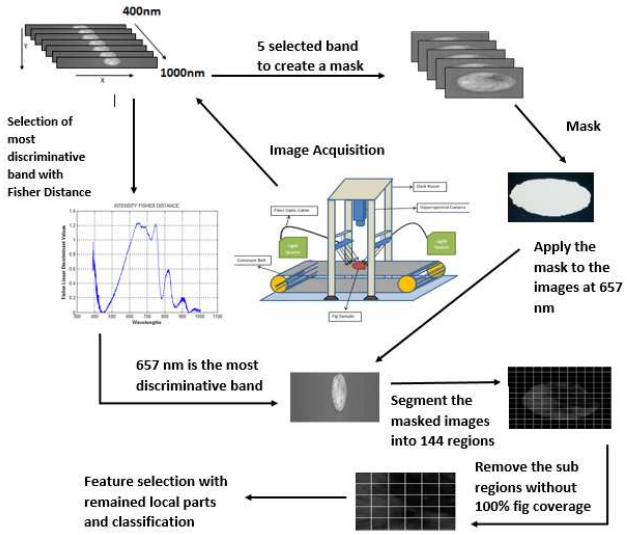

Fig. 2. Flow diagram for detecting contaminated figs.

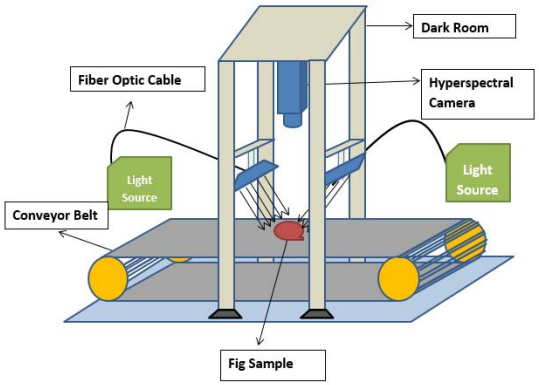

Fig. 3. Hyperspectral imaging system with a conveyor belt and two halogen light sources.

range 400-1000 nm. During image acquisition, the samples are placed on the conveyor belt and the system is operated in the darkened room where only the halogen light sources are used for illumination. The spectral images are then acquired based on SpectralDAQ software.

In the second step, the acquired images are processed to correct the noise generated by the camera using white and dark reference images. A white image was obtained using a white sheet for 20 times (W_1, W_2, ..., W_{20}). The dark image was taken with the lights turned off and the camera covered with its cap 20 times (D_1, D_2, ..., D_{20}). These 20 white and 20 dark images were averaged to obtain W_m and D_m. Then, the relative pixel intensities, $RI(x, y)$, are calculated from the acquired image I using eq. 1

$$RI(x,y) = \frac{I(x,y) - D_m(x,y)}{W_m(x,y) - D_m(x,y)} * 100 \qquad (1)$$

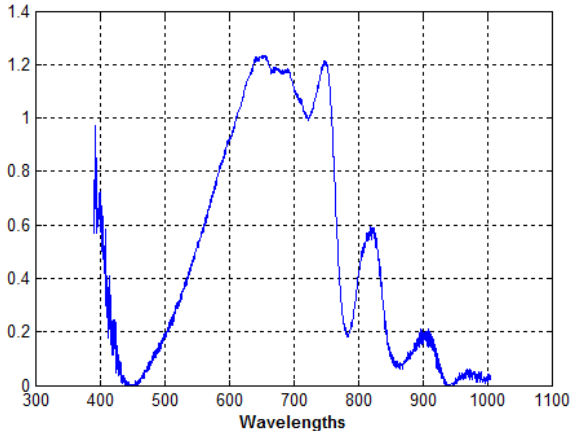

Fig. 4. The Fisher linear discriminant distance calculated from the intensities of the contaminated and uncontaminated classes at each spectral band.

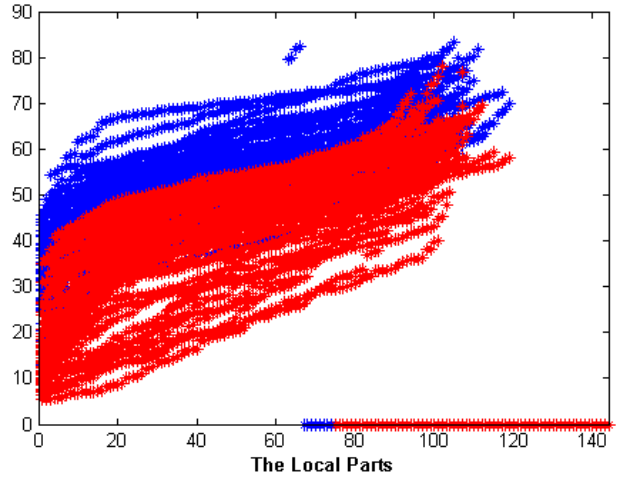

Fig. 5. The intensities of the 144 local parts for each sample.

After line-scan images are acquired, an averaged mask is created from the spectral image at 5 spectral bands from 618.92 nm to 622.03 nm to extract the figs from the background (conveyor belt). These wavelengths give the maximum contrast between the fig and the background.

In the third step, the most informative wavelengths are selected from the hyperspectral bands to avoid computational complexity while keeping the bands with the most discriminative characteristics according to the contaminated and uncontaminated sample classes. For this purpose, Fisher linear discrimination criterion (FLD), which maximizes the ratio of between-class to within-class scatter, is calculated based on the intensities at each wavelength to decide the maximum informative wavelength providing the best discrimination. FLD searches for a vector w that maximizes the separation between the means of positive and negative samples, while minimizing their scattering [13]. By this discrimination method, the highest value of the discrimination between the classes is observed at 657 nm (Fig. 4). Then, the averaged mask is applied to the images at 657 nm.

In the fourth step, each fig image at 657 nm is segmented into equal-sized 144 local squares, due to the fact that the contaminated figs often have low intensity values where contamination is affective (however this region varies in each fig). The squares only with 100% fig coverage are considered for each sample. Since different numbers of these local squares are discarded for each sample, the smallest number of the remained squares is taken as the number of features for all samples (66 out of 144). The average intensity of these 66 squares constitutes 66 features for the classification.

The last step consists of selecting the most significant features among these 66 using forward feature selection (FFS) considering the Mahalanobis distances of the feature subset based on commonly used classifiers Linear Bayes Normal Classifier (LDC), Logistic Linear Classifier (LOGLC), Quadratic Bayes Normal Classifier (QDC), K-Nearest Neighbor Classifier (KNNC with $k = 14$) and Support Vector Classifier (SVC). The FFS starts with identifying the most discriminative feature and iteratively adds the next best feature according to the selection criteria.

3. EXPERIMENTAL RESULTS

In our experiments, 120 samples of dried figs are obtained from the TARIŞ Fig Cooperation, Aydin, Turkey. The samples are labelled by experts into black mold contaminated (60 samples) and uncontaminated samples (60 samples). As explained in Section II, their hyperspectral images are acquired and then extracted fig regions at band 657nm are segmented into 144 subregions. Fig. 5 shows the average intensity values of the 144 subregions for all 120 samples. Note that only 66 of them are used as features.

The most discriminative 30 features were selected among the 66 features at 657 nm, by forward feature selection [14]. Each feature is selected iteratively starting from the most significant to the least significant in terms of accuracies based on five different classifiers. The classification is performed by 10-fold cross validation 10 times and the mean accuracies are shown in Fig. 6 upto 30 most significant features. Table 1 shows the selected accuracies for the number of features between 6 and 23.

While the accuracies are approximately 89% for almost all classifiers for the optimum feature, the accuracies improve to some extent with increased number of features. The greatest accuracies among all classifiers between 9 and 30 features are obtained by LDC, with a maximum accuracy of 93.58% with the first 23 selected features. The greatest accuracies for five classifiers are as follows: 93.58% with 23 features for LDC, 90.50% with 6 features for QDC, 91.50% with 17 fea-

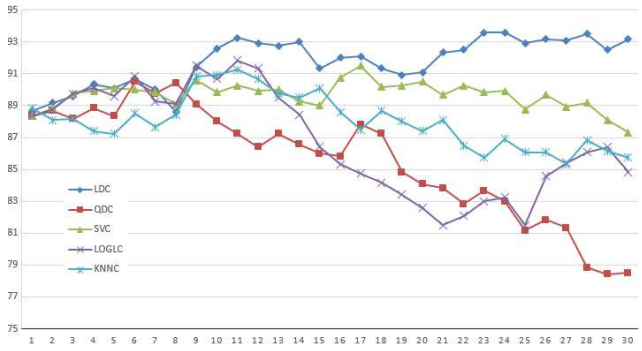

Fig. 6. The accuracies (averaged over 10 times 10-fold cross validation) with respect to the number of features.

Table 1. Accuracies for the number of features (6 to 23).

# features	LDC	QDC	SVC	LOGLC	KNNC
6	90.67	90.50	90.00	90.83	88.50
7	90.00	89.75	89.83	89.25	87.67
8	88.58	90.42	89.08	89.08	88.42
9	91.42	89.08	90.58	91.50	90.83
10	92.58	88.00	89.83	90.67	90.92
11	93.25	87.25	90.25	91.83	91.25
12	92.92	86.42	89.92	91.33	90.67
13	92.75	87.25	90.00	89.50	89.75
14	93.00	86.58	89.25	88.42	89.50
15	91.33	86.00	89.00	86.42	90.08
16	92.00	85.83	90.75	85.33	88.58
17	92.08	87.83	91.50	84.75	87.50
18	91.33	87.25	90.17	84.17	88.67
19	90.92	84.83	90.25	83.42	88.00
20	91.08	84.08	90.50	82.58	87.42
21	92.33	83.83	89.67	81.50	88.08
22	92.50	82.83	90.25	82.08	86.50
23	93.58	83.67	89.83	83.00	85.75

tures for SVC, 91.50% with 9 features for LOGLC, 91.25% with 11 features for KNNC respectively. These preliminary classification accuracies are promising for an operational detection system. However, currently they are computationally expensive for an online system.

4. CONCLUSION

Hyperspectral imaging systems are increasingly becoming popular for automatic food quality assessment. We proposed such a system for classification of dried figs into black mold contaminated and uncontaminated ones. Our initial setup consisted of selecting the best discriminative spectral band and spatial assessment on this single band. We achieved an accuracy of 93.58% using twenty-three most discriminative features (local averages of intensities) and linear discriminant classifier. This accuracy is already promising and very close to the manual assessment which requires a heavy workload and expensive process. Despite this successful classification, a detailed feature selection considering all available spectral bands, their spatial correlation and the transmittance values of figs at various bands will help better distinguish the contaminated figs, to achieve an operational system.Alternatively, transmittance spectroscopy can help better identification of infected figs[15].

5. REFERENCES

[1] Di Wu and Da-Wen Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review Part II: Applications," *Innovative Food Science & Emerging Technologies*, vol. 19, pp. 15–28, 2013.

[2] Hui Huang, Li Liu, and Michael O Ngadi, "Recent developments in hyperspectral imaging for assessment of food quality and safety.," *Sensors (Basel, Switzerland)*, vol. 14, pp. 7248–76, 2014.

[3] Gamal ElMasry and Da-wen Sun, "Principles of Hyperspectral Imaging Technology," in *Hyperspectral Imaging for Food Quality Analysis and Control*, chapter Principles, pp. 3–43. Elsevier, 2010.

[4] A. A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias, "Hyperspectral imaging - an emerging process analytical tool for food quality and safety control," *Trends in Food Science and Technology*, vol. 18, pp. 590–598, 2007.

[5] Wenwen Kong, Fei Liu, Chu Zhang, Yidan Bao, Jiajia Yu, and Yong He, "Fast detection of peroxidase (POD) activity in tomato leaves which infected with Botrytis cinerea using hyperspectral imaging," *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, vol. 118, pp. 498–502, 2014.

[6] Jiangbo Li, Xiuqin Rao, and Yibin Ying, "Detection of common defects on oranges using hyperspectral reflectance imaging," *Computers and Electronics in Agriculture*, vol. 78, no. 1, pp. 38–48, 2011.

[7] Jianwei Qin and Renfu Lu, "Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique," *Postharvest Biology and Technology*, vol. 49, no. 3, pp. 355–365, 2008.

[8] R Lu, "Detection of bruises on apples using near-infrared hyperspectral imaging," *Transaction of the American Society of Agricultural Engineers*, vol. 46, no. 2, pp. 1–8, 2003.

[9] Douglas F. Barbin, Gamal Elmasry, Da Wen Sun, Paul Allen, and Noha Morsy, "Non-destructive assessment of microbial contamination in porcine meat using NIR hyperspectral imaging," *Innovative Food Science and Emerging Technologies*, vol. 17, pp. 180–191, 2013.

[10] Gabriel A. Leiva-Valenzuela, Renfu Lu, and José Miguel Aguilera, "Assessment of internal quality of blueberries using hyperspectral transmittance and reflectance images with whole spectra or selected wavelengths," *Innovative Food Science and Emerging Technologies*, vol. 24, pp. 2–13, 2014.

[11] Renfu Lu and Diwan P. Ariana, "Detection of fruit fly infestation in pickling cucumbers using a hyperspectral reflectance/transmittance imaging system," *Postharvest Biology and Technology*, vol. 81, pp. 44–50, 2013.

[12] M. A. Doster, "Aspergillus Species and Mycotoxins in Figs from California Orchards," *Plant Disease*, vol. 80, pp. 484–489, 1996.

[13] Christopher M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[14] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, "The Elements of Statistical Learning," *Elements*, vol. 1, pp. 337–387, 2009.

[15] Efkan Durmuş, Ahmet.S. Bilgi, Gizem Ortaç, Habil Kalkan, and Kadim Taşdemir, Eds., *Detection of Black Mold Infected Figs by Using Transmittance Spectroscopy*. in Proc. Workshop on Hyperspectral Image and Signal Processing WHISPERS'2015,2-5 June, Tokyo, Japan, 2015.