The final publication is available at

http://dx.doi.org/10.1007/s11947-011-0737-x

Additional Information

Selection of Optimal Wavelength Features for Decay Detection in Citrus Fruit Using the ROC Curve and Neural Networks

Delia Lorente · Nuria Aleixos · Juan Gómez-Sanchis · Sergio Cubero · Jose Blasco

Abstract Early automatic detection of fungal infections in post-harvest citrus fruits is especially important for the citrus industry because only a few infected fruits can spread the infection to a whole batch during operations such as storage or exportation, thus causing great economic losses. Nowadays, this detection is carried out manually by trained workers illuminating the fruit with dangerous ultraviolet lighting. The use of hyperspectral imaging systems makes it possible to advance in the development of systems capable of carrying out this detection process automatically. However, these systems present the disadvantage of generating a huge amount of data, which must be selected in order to achieve a result that is useful to the sector. This work proposes a methodology to select features in multi-class classification problems using the receiver operating characteristic curve, in order to detect rottenness in citrus fruits by means of hyperspectral images. The classifier used is a multilayer perceptron, trained with a new learning algorithm called extreme learning machine. The results are obtained using images of mandarins with the pixels labelled in five different classes: two kinds of sound skin, two kinds of decay and scars. This method yields a reduced set of features and a classification success rate of around 89%.

D. Lorente · S. Cubero · J. Blasco (✉)
Centro de Agroingeniería,
Instituto Valenciano de Investigaciones Agrarias (IVIA),
Cra. Moncada-Náquera km 5,
46113 Moncada, Valencia, Spain
e-mail: blasco_josiva@gva.es

N. Aleixos
Instituto Interuniversitario de Investigación en Bioingeniería y
Tecnología Orientada al Ser Humano,
Universitat Politècnica de València,
Camino de Vera s/n,
46022 Valencia, Spain

J. Gómez-Sanchis
Intelligent Data Analysis Laboratory, IDAL,
Electronic Engineering Department, Universitat de València,
Dr. Moliner 50,
46100 Burjassot, Valencia, Spain

Introduction

Decay pathogens can enter fruit through wounds sustained during harvesting. This implies that the pathogen is already in the fruit before any treatment is applied in post-harvest (Obagwu and Korsten 2003). Early detection of fungal infections in citrus fruits is especially important in packing-houses because a very small number of infected fruits can spread the infection to a whole batch, thus causing great economic losses and affecting further operations, such as storage and transport. The most important post-harvest damage in citrus packinghouses is caused by *Penicillium* sp. fungi (Eckert and Eaks 1989). Nowadays, the detection of rotten fruit on citrus packing lines is carried out visually under dangerous ultraviolet (UV) illumination, and decay fruits are removed manually. This procedure, however, may be harmful for operators and operationally inefficient, since they must work in shifts of just a few hours. This rate of staff rotation affects the assessment of the quality. A possible solution arises from the use of automatic machine vision systems.

Computer vision has become widely used to automate the inspection of all different types of food commodities like meat (Du and Sun 2009), fish (Quevedo and Aguilera

2010; Quevedo et al. 2010), bakery products (Farrera-Rebollo et al. 2011), grains (Manickavasagan et al. 2010) or fruits (Karimi et al. 2009). In most cases, its use is aimed at the inspection of external features related with quality, such as size, shape, colour or the presence of damage (Cubero et al. 2011). The use of technology based on colour cameras for the detection of external damage of citrus is currently under research. Kim et al. (2009) used colour texture features based on HSI and colour co-occurrence method to detect peel diseases in grapefruit. López-García et al. (2010) used multivariate image analysis with the same objective in citrus fruits. However, some defects, like decay or freeze damage, are very difficult to detect using standard artificial vision systems since they are hardly visible to the human eye and, consequently, by standard red–green–blue (RGB) cameras Blasco et al. (2007). Therefore, different technologies have to be incorporated, such as the use of UV-induced fluorescence (Slaughter et al. 2008; Obenland et al. 2009). In an attempt to automate the current manual tasks of detection of decay, Blanc et al. (2009, 2010) patented an automatic machine for decay detection using UV illumination, and Kurita et al. (2009) developed an inspection system based on simultaneous visible and UV illumination using light-emitting diodes. However, it would be desirable to avoid the use of UV radiation in these tasks which could be achieved by finding out particular wavelengths in the visible or near-infrared (NIR) part of the electromagnetic spectrum.

Images acquired in visible and NIR simultaneously were used to detect different types of damage in citrus fruits by Aleixos et al. (2002) and more recently by Blanc et al. (2009), who attempted to detect common external defects and diseases, including decay, by combining NIR, visible and also UV-induced fluorescence. In this sense, the recent introduction of hyperspectral sensors for the inspection of food (Sun 2010) makes it possible to carry out a more precise analysis of the problem by acquiring images for specific ranges of wavelengths to detect features non-visible features or to select particular sets of some wavelengths related to important physical properties, as indicated in the review of Lorente et al. (2011).

Using spectroscopy, Gaffney (1973) found that different external defects on citrus fruits have different spectral signatures, stated later in the review of Magwaza et al. (2011), which can lead to the selection of certain sets of wavelengths that facilitate the detection of particularly dangerous defects such as canker (Balasundaram et al. 2009). However, in real life, it is not enough just to distinguish between fruit affected by serious diseases and sound fruit. It is important to develop systems capable of separating also produce affected by scars on the rind, or other external defects that only downgrade the quality of the fruit but do not spread among other fruits and do not prevent its marketing in domestic markets Blasco et al. (2009). If they are not taken into account, these cosmetic defects may be confused with the dangerous by an automatic system. Qin et al. (2009) used a hyperspectral system with sensitivity in the range 450–930 nm to detect different kinds of damage that affect the skin of citrus, with particular attention being paid to the detection of canker from other common defects. However, one of the main problems of these systems is the huge amount of data generated (Gómez-Sanchis et al. 2008a).

While a standard RGB image is composed of three images corresponding to the red, green and blue bands, a hyperspectral image consists of a set of monochromatic, narrow-band images that increases the complexity of the analysis and requires more computing time to analyse them with an automatic system, which prevent its use in real-time in-line inspection system. For this reason, it is very important to select only those bands with the most relevant information, while discarding those that do not contribute in any significant way to improve the results. With the aim of detecting different defects on the skin of oranges using a hyperspectral system, Li et al. (2011) used principal component analysis (PCA) to select two sets of six and three optimal wavelengths and later applied PCA and band ratios to detect the defects in these multispectral images.

Generally, statistical methods to reduce dimensionality and select features can be divided into wrapper and filter methods (Guyon and Elisseeff 2003). Filter methods use an indirect measure of the quality of the selected features (e.g. by evaluating the correlation function between each input feature and the dependent variable—class—of the classification problem), obtaining a faster convergence of the selection algorithm. On the other hand, the selection criteria used by wrapper methods are the goodness of fit between the inputs and the output provided by the learning machine under consideration, like, for example, a neural network. Within these methods, a traditional measure for evaluating classifiers is the classification success rate. However, a more suitable way of measuring the quality of a classifier is the area under the receiver operating characteristic (ROC) curve, which is the measure used in the feature selection method proposed in this work. Basic concepts related to classification models are first reviewed for a better understanding of the ROC curve as feature selection method. The ROC curve is a graphical plot of the true positive rate vs. false positive rate for a binary classifier, as its discrimination threshold is varied, this value being defined as that from which a positive class prediction is made (Fawcett 2006). The area under the ROC curve (AUC) is used as a global measure of classifier performance that is invariant to the classifier discrimination threshold and the class distribution (Bradley 1997). Maximum classification accuracy corresponds to an AUC value

of 1, while a random guess separation involves a minimum AUC value of 0.5.

With regard to classification methods, because of their flexibility and the possibility of working with unstructured and complex data like that obtained from biological products, artificial neural networks (ANN) have been applied in almost every aspect of food science, and it is a useful tool for performing food safety and quality analyses. For instance, a combination of principal components analysis and ANN was used by Bennedsen et al. (2007) to detect surface defects on apples. Unay and Gosselin (2006) used a multilayer perceptron (MLP) as a promising technique for segmenting surface defects on apples. Ariana et al. (2006) presented an integrated approach using multispectral imaging in reflectance and fluorescence modes to acquire images of three varieties using two ANN-based classification schemes (binary and multi-class). In the case of citrus fruits, Kondo et al. (2000) used, among other methods, ANN to detect some external and internal features in oranges while Gómez-Sanchis et al. (2012) used minimum redundancy maximum relevance as feature selection method and MLP for pixel classification to detect rottenness in mandarins.

This paper advances in the automatic detection of a dangerous post-harvest disease of citrus fruits, such as fungal decay, and to distinguish fruit with symptoms of decay from sound fruit and affected by minor defects. A feature selection methodology that expands the use of the ROC curve to multi-class classification problems is proposed. This methodology has been applied to the selection of an optimal set of features that are effective in the detection of common defects and decay in citrus fruits using hyperspectral images.

In particular, we have used computer vision for detection of two dangerous types of decay caused by *Penicillium digitatum* Sacc (green mould) and *Penicillium italicum* Wehmer (blue mould) because these pathogens occur in almost all regions of the world where citrus is grown and cause serious post-harvest losses annually (Palou et al. 2001). Furthermore, in order to explore the possibilities of the ROC method as a technique for selecting important wavelengths in fruit inspection, we used an ANN-based classifier trained with a new learning algorithm called extreme learning machine (ELM; Huang et al. 2006).

Feature Selection Methodology

Imaging System

In this work, a hyperspectral vision system based on liquid crystal tunable filters (LCTF) was employed. The set of monochrome images acquired by this system makes up a hyperspectral image from which spatial as well as spectral information can be obtained about the scene. A hyperspectral image can be interpreted as a hypercube, in which two dimensions are spatial (pixels) and the third is the spectrum of each pixel. The system consists of a monochrome camera (CoolSNAP ES, Photometrics) with a high level of sensitivity between 320 and 1,020 nm. It was set to acquire 551×551 pixel images with a resolution of 3.75 pixels/mm. The camera transfers the images to a computer by means of a proprietary frame grabber based on PCI technology. The computer employed is based on a Pentium 4 processor with 1 Gb of random access memory. A lens capable of providing a uniform focus between 400 and 1,000 nm was chosen for use with the system (Xenoplan 1.4/17MM, Schneider).

Two LCTF were used, one sensitive to the visible between 400 and 720 nm (Varispec VIS07, CRI Inc) and one sensitive to NIR in the 650- to 1,100-nm range (Varispec NIR07, CRI Inc). Each fruit was illuminated individually by indirect light from 12 halogen lamps (20 W) inside an aluminium hemispherical diffuser in order to provide good spectral efficiency in the visible and NIR. The lamps were powered by a stabilised power supply (12 V/DC 350 W). Because the sum of efficiencies of the filter, camera and illumination system is different across the selected wavelengths, the acquisition software was programmed to correct the integration time for each particular band that is acquired. Hence, these differences in the efficiency of the filter for each band are offset by calculating the particular integration time for each image in each wavelength using a white reference, so that the spectral response of the system is flat over the whole spectral range.

The filters were placed just in front of the camera lens. One of the main problems arose when it came to changing between visible and infrared filters, since the camera could move when handling the filters, which made it difficult to acquire the exactly same scene with both filters. This problem was solved by designing and installing a system to hold and guide the filters. The two filters were fitted to the support and move on a sliding mechanism, thus allowing each filter to be set in the right place without handling the camera. The arrangement of the image acquisition system inside the inspection chamber is shown in Figure 1.

Fruit Used in the Experiments

The experiments were carried out using mandarins cv. Clemenules (*Citrus clementina* Hort. ex Tanaka) with two kinds of defects: (1) minor defects represented by external scars affecting only the appearance of the fruit and (2) serious diseases that can spread to other fruits caused by

Fig. 1 Scheme of the image acquisition system showing the arrangement of the visible and near-infrared liquid crystal tunable filters

two different fungi; *P. digitatum* and *P. italicum*. The fruits affected by the first type of defects were chosen at random from the packing line of a trading company. On the other hand, damage produced by fungi was caused artificially in sound fruits using an inoculation of spores.

A total of 240 fruits were used: 60 sound fruits, 60 presenting external scars, 60 inoculated with spores of *P. digitatum* and 60 inoculated with spores of *P. italicum*. The inoculation was performed using a suspension of spores with a concentration of 10^6 spores/ml for both fungi, which is sufficient to cause infestation in laboratory conditions (Palou et al. 2001). From the point of view of the post-harvest, it is probably not important to differentiate between both types of decay. However, in this paper, this distinction has been made to test the potential of this method to discriminate between defects that are virtually identical in their early stages to the naked human eye. The fruits were stored for 3 days in a controlled environment at 25 °C and 99% relative humidity. After this period, all the inoculated fruits presented a characteristic patch of rottenness with a diameter between 10 and 35 mm. While rind scars are clearly visible, the colour of rotten skin is similar to the colour of the sound skin around it, thus making it difficult for a human inspector to detect it.

The images were acquired by placing the fruit manually in the inspection chamber and then presenting the damage to the camera. A total of 240 hyperspectral images were acquired from 460 to 1,020 nm, with a spectral resolution of 10 nm. The hyperspectral image was therefore composed of 57 monochrome images of each fruit, which gives a total number of 13,680 monochrome images. The analysis of images started by correcting the effects of illumination on spherical fruits following the methodology described in Gómez-Sanchis et al. (2008b). Then, in order to separate the fruit from the background in the image, the hyperspectral images were pre-processed using masking. The mask was created by thresholding the fruit image at

650 nm, since images at this wavelength provided the best contrast between fruit and background.

Figure 2 shows the RGB images of four fruits corresponding to a sound fruit, a fruit with scars on the rind, a fruit infected by *P. digitatum* and a fruit infected by *P. italicum*, respectively, from top to bottom. The adjoining columns show example images of the same fruits acquired at 530, 640, 740 and 910 nm, respectively. These images were chosen at different wavelengths just to have an overall impression of what could be seen in hyperspectral images but not necessarily used in the experiments. In the RGB images, the damage caused by fungi is hardly visible to the naked human eye.

Labelled Set

In supervised classification, there is a set of n labelled samples, $\{x_i, y_i\}_{i=1..n}$, where x_i represents the m-dimensional feature vector for the ith pixel with label y_i. Here, m represents the spectral bands and spectral indexes, and y defines the universe of all possible labelled classes in the image. In this work, the supervised nature of the problem presented here required the construction of a labelled data set, consisting of $m=74$ features associated to each pixel, specifically 57 purely spectral variables (reflectance level of the pixel for each acquired band) and 17 spectral indexes calculated by combining several reflectance values, as shown in Table 1. The spectral indexes were used to know if any of them could improve the decay detection in comparison to the use of only purely spectral variables. In order to build this labelled set, $n=143,095$ pixels were selected manually, and then a human expert assigned them to one of the five classes considered in this work: green sound skin (GS), orange sound skin (OS), defective skin by scars (SC), decay caused by *P. digitatum* (PD) and decay caused by *P. italicum* (PI). Each sample pattern is therefore composed of 74 features and a class label. A background class was not included, since the background pixels were segmented earlier in the pre-processing step.

The labelled set was divided into a training set of 35,774 samples (25% of the total), a validation set of 35,774 samples (25% of the total) and a test set of 71,547 samples (50% of the total). The first two sets were used to build the proposed statistical methods of feature selection and classification and the third one to evaluate classifier performance. The choice of a huge number of pixels in the test set was made in order to check the generalisation capability of the models.

Feature Selection

The feature selection methodology proposed to expand the use of the ROC curve to multi-class classification

problems consists of two parts: (1) obtaining a ranking of features ordered according to the discriminant relevance of the features and (2) the choice of an optimal number of features from the feature ranking. Both this feature selection method and the classification procedure used in this work were implemented using Matlab 7.9 (Mathworks, Inc.).

Obtaining a Feature Ranking

The first step consists in obtaining a feature ranking for each class. The ROC curve is intended for binary classification problems. However, in this work, problems with more than two classes are considered. Therefore, the *one* vs. *all* (OVA) approach is employed to obtain a feature ranking for each class, which maximises the separation between that class and the others. The OVA structure consists in assuming that the problem has only two classes: a class from which the ranking is obtained and another class grouping the remaining classes (Rifkin and Klautau 2004). In order to obtain these partial rankings, several steps were followed for each class, these steps being similar to the ones used by Serrano et al. (2010) in binary classification problems; the classifier is trained using all the features, taking into account the OVA structure, that is, considering a classification problem with only two classes. Then, the area under the ROC curve is obtained for the classification model using all features (AUC$_0$). The following parameters are obtained for each input feature x_i:

- Area under the ROC curve for the classifier without taking into account the effect of feature x_i (AUC$_i$). For this purpose, when using the classifier, the feature x_i is assumed to be constant for every sample, $x_i=0$.

- Discriminant relevance of feature x_i (DR$_i$), which is defined as the difference between the area under the ROC curve of the classifier using all the features (AUC$_0$) and the area without taking into account the effect of feature x_i (AUC$_i$). This parameter indicates the importance of a feature for the discrimination process carried out by the classifier, considering that the higher the discriminant relevance of a feature is, the more discriminatory that feature will be.

- A z statistic of feature x_i (z_i) is calculated from the discriminant relevance of feature x_i (DR$_i$), as shown in Eq. 1:

$$z_i = \frac{\text{AUC}_0 - \text{AUC}_i}{\sqrt{\text{SE}_0^2 + \text{SE}_i^2 + 2 \cdot \rho \cdot \text{SE}_0 \cdot \text{SE}_i}} \tag{1}$$

where SE$_0$ and SE$_i$ are the standard errors of AUC$_0$ and AUC$_i$, respectively, and ρ is the correlation between AUC$_0$ and AUC$_i$. In this work, a feature is considered to be relevant for the problem when its corresponding z value exceeds 95%, this level being chosen empirically.

Features in each ranking are ordered according to the contribution each of them makes to the discriminant capability of the classification process, and the input features with the highest z values are considered the most discriminatory features.

The second step consists in obtaining a global feature ranking. After obtaining the partial rankings corresponding to each class, the next step is to perform a single global ranking that includes all the classes. The z values corresponding to the rankings for each class are combined by means of their weighted mean (Eq. 2), which assigns a weight to each class in proportion to its relative importance

Table 1 Spectral indexes used in this work as input features

Vegetative index	Estimation	Parameter values used in this work	Reference
NDVI	$NDVI = \frac{R_{NIR} - R_{RED}}{R_{NIR} + R_{RED}}$	λ_{NIR}=800 nm λ_{RED}=640 nm	Tucker (1979)
Green NDVI. Version I	$Green\ NDVI\ I = \frac{R_{GREEN} - R_{RED}}{R_{GREEN} + R_{RED}}$	λ_{GREEN}=550 nm λ_{RED}=640 nm	Yang et al. (2007)
Green NDVI. Version II	$Green\ NDVI\ II = \frac{R_{NIR} - R_{GREEN}}{R_{NIR} + R_{GREEN}}$	λ_{NIR}=800 nm λ_{GREEN}=550 nm	Gitelson et al. (1996)
WBI	$WBI = \frac{R_{950\,nm}}{R_{900\,nm}}$		Xu et al. (2007)
SAVI	$SAVI = \frac{(R_{NIR} - R_{RED})(1+L)}{R_{NIR} + R_{RED} + L}$	λ_{NIR}=800 nm λ_{RED}=640 nm L=0.5	Yang et al. (2007)
PRI	$PRI = \frac{R_{531\,nm} - R_{570\,nm}}{R_{531\,nm} + R_{570\,nm}}$	$\lambda_{531\,nm} \approx 500$ nm	Huang et al. (2007)
RVSI	$RVSI = \frac{R_{714\,nm} + R_{752\,nm}}{2 - R_{733\,nm}}$	$\lambda_{714\,nm} \approx 710$ nm $\lambda_{752\,nm} \approx 750$ nm $\lambda_{733\,nm} \approx 730$ nm	Naidu et al. (2009)
MCARI	$MCARI = [(R_{700\,nm} - R_{670\,nm}) - 0.2(R_{700\,nm} - R_{550\,nm})] \times \frac{R_{700\,nm}}{R_{670\,nm}}$		Naidu et al. (2009)
VARI	$VARI = \frac{R_{GREEN} - R_{RED}}{R_{GREEN} + R_{RED} - R_{BLUE}}$	λ_{GREEN}=550 nm λ_{RED}=640 nm λ_{BLUE}=480 nm	Naidu et al. (2009)
WI	$WI = \frac{R_{900\,nm}}{R_{970\,nm}}$		Naidu et al. (2009)
TCARI	$TCARI = 3 \times [(R_{700\,nm} - R_{670\,nm}) - 0.2(R_{700\,nm} - R_{550\,nm})] \times \frac{R_{700\,nm}}{R_{670\,nm}}$		Haboudane et al. (2002)
OSAVI	$OSAVI = \frac{(R_{NIR} - R_{RED})(1+0.16)}{R_{NIR} + R_{RED} + 0.16}$	λ_{NIR}=800 nm λ_{RED}=640 nm	Rondeaux et al. (1996)
CCI	$CCI = \frac{a \times 1,000}{L \times b}$	a, b and L are the coordinates of the CIELAB colour space	Jiménez-Cuesta et al. (1981)
Other indexes	$R_{NIR} - R_{RED}$	λ_{NIR}=800 nm λ_{RED}=640 nm	Yang et al. (2007)
	$\frac{R_{RED}}{R_{NIR}}$	λ_{NIR}=800 nm λ_{RED}=640 nm	Yang et al. (2007)
	$\frac{R_{GREEN}}{R_{RED}}$	λ_{GREEN}=550 nm λ_{RED}=640 nm	Yang et al. (2007)
	$\frac{R_{NIR}}{R_{RED}}$	λ_{NIR}=800 nm λ_{RED}=640 nm	Yang et al. (2007)

R_{λ} is the reflectance value at band λ

NDVI normalised difference vegetation index, *WBI* water band index, *SAVI* soil-adjusted vegetation index, *PRI* photochemical reflectance index, *RVSI* red-edge vegetation stress index, *MCARI* modified chlorophyll absorption in reflectance index, *VARI* visible atmospherically resistant index, *WI* water index, *TCARI* transformed chlorophyll absorption in reflectance index, *OSAVI* optimised soil-adjusted vegetation index, *CCI* citrus colour index

in the classification problem. Thus, the global relevance of each feature is obtained, and then each input feature is ranked according to its global relevance. The ranking thus obtained maximises the global separation among all the classes.

$$\bar{z}_i = \frac{\sum_{k=1}^{N} z_{ik} \cdot w_k}{\sum_{k=1}^{N} w_k} \tag{2}$$

where $\bar{z}_i$ is the global relevance of feature x_i, N is the number of different classes, z_{ik} is the z value of feature x_i from the partial ranking for the kth class and w_k is the weight for the kth class.

Choice of an Optimal Number of Features

In this stage, a minimum number of features leading to a saturation trend in the success rate of classification are chosen. The following steps are required to do this:

The initial step is to obtain the evolution of the success rate of classification as a function of the number of features. For this purpose, the classifier is trained using

the first feature of the global ranking and its success rate is evaluated, this process is then repeated including the next feature of the ranking and so on, until all the features are employed sequentially. Then, the first number of input features n satisfying the two conditions in Eqs. 3 and 4 is chosen, where $success_n$ is the success rate of classification using n features, $success_{n+1}$ the success rate with $n+1$ input features and so on.

$$success_{n+1} - success_n \leq 1\% \qquad (3)$$

$$success_{n+2} - success_{n+1} \leq 1\% \qquad (4)$$

Classifier

The classifier used to explore the possibilities of the proposed feature selection methodology is a MLP with a single hidden layer, which is the simplest kind of ANN. However, the feature selection procedure is independent of the chosen classification method. ANN is considered to be a commonly used pattern recognition tool in hyperspectral image processing because it is capable of handling a large amount of heterogeneous data with considerable flexibility and has non-linear properties (Plaza et al. 2009).

The most popular ANN is the MLP, which is a feed-forward ANN model that maps sets of input data onto a set of appropriate output and consists of multiple layers of nodes (neurons) in a directed graph that is fully connected from one layer to the next. In particular, the MLP used in this work has an input layer, a single hidden layer and an output layer. MLP can use a large variety of learning techniques, the most popular being backpropagation, which is a supervised learning method based on gradient descent in error that propagates classification errors back through the network and uses those errors to update parameters (Shih 2010). In these classical learning methods, the parameters of the ANN are normally tuned iteratively and thus entail several disadvantages, such a high degree of slowness and convergence to local minima. In order to avoid these problems, the MLP

used in this work was trained using ELM, which is a new learning algorithm that determines the ANN parameters (not the optimal architecture) analytically in a faster way instead of tuning them iteratively. This increase of speed in the learning algorithm is very important in order to search the optimal features in our particular feature selection problem using ROC curve. Moreover, this learning algorithm for feed-forward neural networks with a single hidden layer, like an MLP, provides good generalisation performance, as well as an extremely fast learning speed (Huang et al. 2006).

Considering a set of n patterns, $\{x_i, t_i\}_{i=1..n}$ and M nodes in the hidden layer, the MLP output for the ith sample is given by Eq. 5, which is obtained in a straightforward way taking into account the structure of an artificial neuron, as well as the MLP structure (Fig. 3).

$$y_i = \sum_{j=1}^{M} g(\mathbf{w}_j \cdot \mathbf{x}_i) \cdot \beta_j \qquad (5)$$

where $\mathbf{w}_j$ is the weight vector connecting the jth hidden node and the input nodes, β_j is the weight vector connecting the jth hidden node and the output nodes and g is an activation function applied to the scalar product of the input vector and the hidden layer weights.

Equation 6 can be written compactly in matrix notation as $\mathbf{y} = \mathbf{G} \cdot \boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is the weight vector of the output layer and $\mathbf{G}$ is given by:

$$G = \begin{pmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_n) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_n) \end{pmatrix} \qquad (6)$$

ELM proposes a random choice of the weights of the hidden layer, $\mathbf{w}_j$, thus making it necessary only to determine the weights of the output layer, $\boldsymbol{\beta}$, analytically through simple generalised inverse operation of the matrix $\mathbf{G}$ according to the Eq. 7:

$$\beta = G^{\dagger} \cdot t \qquad (7)$$

where $G^{\dagger} = (G^T \cdot G)^{-1} \cdot G^T$ is the Moore–Penrose generalised inverse of matrix $\mathbf{G}$ (Rao and Mitra 1972), $\mathbf{G}^T$ being the transpose of matrix $\mathbf{G}$.

Figure 3 Structures of a multilayer perceptron with a single hidden layer (*left*) and an example of artificial neuron (*right*)

An important issue in practical applications of ELM is how to obtain an optimal number of the hidden nodes in the network architecture in order to achieve a good generalisation performance when training a neural network. The methodology used to select the optimum number of hidden neurons was to estimate the classification success rate for several models, obtained by varying the number of neurons in the hidden layer (Huang et al. 2006). In a first step, architectures with a variable number of hidden neurons from 25 to 1,025 in increments of 100 elements were tested in order to obtain the range of the architectures that fit correctly the data maintaining the generalisation capabilities of the model. These limits were set because networks that are too small cannot model the data properly, while networks that are too large may lead to overfitting (Prechelt 1996). Attending the curve of success rate, the optimum range was selected between 75 and 225 neurons. In a second step, architectures using 75 from 225 neurons were tested selecting finally a MLP that used $M=125$ neurons in the hidden layer and the sigmoid function as the activation function (g). The classification success rate for the model with 125 hidden neurons was 91.4%, while the success rate for the model with 1,025 neurons was about 95.6%, thus improving by only 4% while the training time burst.

Figure 4 The z statistic of the 74 input features for each of the five classes: defects by scars on the rind, green sound skin, orange sound skin, decay caused by *P. digitatum* and decay caused by *P. italicum*. *Horizontal solid lines* indicate the limit at the 95% level

Table 2 Selected features and their correspondence with the spectral bands or for approach I

Input feature	Spectral index or reflectance value
24	Reflectance at 690 nm
10	Reflectance at 550 nm
58	NDVI
27	Reflectance at 720 nm
16	Reflectance at 610 nm
50	Reflectance at 950 nm

NDVI normalised difference vegetation index

Approaches to the Problem of Decay Detection

This work considers three different approaches to feature selection in the problem of the detection of decay in mandarins, depending on the number of classes involved in the problem and the weight or importance of each class. The typical problem involves the five classes described in the labelled set, all of them having equal importance or weight.

The first approach considers five different classes of similar importance in the classification problem. Therefore, when obtaining the global relevance of each feature, the weights of all the classes were considered to be equal. The aim of this approach is to know the behaviour of the method by considering a quality classification of the fruit, which separates sound fruits from those that only contain cosmetic effects that degrade the appearance and from dangerous infections. However, it is reasonable to assume that in the real world, the classes belonging to decaying skin should be more important for the problem which is the detection of decay.

Therefore, the approach II rests on the idea that the problem has five classes of different importance in the classification. To know the behaviour of the proposed method to enhance the detection of these most important cases, empirical weights were assigned to the classes in Eq. 2, more importance being given to decay classes ($w_{PD}=w_{PI}=15$), medium importance was given to the scar class ($w_{SC}=5$) and less to sound classes ($w_{GS}=w_{OS}=1$).

Moreover, decay is the disease whose detection is of most importance and which has still not been solved by automatic systems. Hence, since the actual aim of a potential inspection system would be to detect decay, it is also important to study the potential of the detection of just infected fruit, which leads to a binary problem: the separation between infected or not infected fruit (approach III). Two classes were defined:

- Decay. This class includes the two kinds of decay presented in this work: infection caused by *P. digitatum* and by *P. italicum*

- Not decay. This class groups the remaining classes: green sound skin, orange sound skin and scars

Results and Discussion

Feature Selection

Figure 4 shows the z statistic obtained for the 74 input features for each of the five classes. This statistic gives the same information as the variation in *AUC*. In addition, it makes it possible to study whether an input feature is discriminant or not.

After obtaining the z values of the input features for each class, the global relevance of each feature was computed for approaches I and II, considering five classes of similar importance and five classes of different importance, respectively. The resulting optimal number of features according to the proposed mathematical criterion, shown in Eqs. 3 and 4, is six for the first approaches and seven for the second. Table 2 shows the set of selected features for the first approach, as well as the correspondence between these features and the spectral indexes or reflectance values associated to them. Similarly, Table 3 shows the set of selected features, ordered according to their importance in the classification problem, for the second approach and the correspondence between the selected input features and the spectral indexes or reflectance values.

When comparing Tables 2 and 3, it can be noticed that most of the input features are coincident in both sets, except feature 16 for approach I and features 74 and 22 in the case of approach II. This is due to the fact that these two features are really important for the detection of pixels belonging to the classes of decay, the highest weights being achieved when the global value of z is obtained, as can be straightforwardly seen from the z values for the *P. italicum* class in Fig. 4.

Furthermore, feature 16 is not selected for the second approach, while in the first approach it is. This is due to the

Table 3 Selected features and their correspondence with the spectral bands or indexes for approach II

Input feature	Spectral index or reflectance value
24	Reflectance at 690 nm
58	NDVI
27	Reflectance at 720 nm
50	Reflectance at 950 nm
10	Reflectance at 550 nm
74	CCI
22	Reflectance at 670 nm

NDVI normalised difference vegetation index, *CCI* citrus colour index

Table 4 Selected features and their correspondence with the spectral indexes or reflectance values for approach III

Input feature	Spectral index or reflectance value
23	Reflectance at 680 nm
60	Green NDVI, version II
28	Reflectance at 730 nm
15	Reflectance at 600 nm

NDVI normalised difference vegetation index

fact that, although this feature has a high level of importance for the classification of pixels belonging to the orange skin class, as shown in Fig. 4, it is considered of low importance when obtaining the global relevance in the second approach. In addition, a general conclusion drawn from analysing the results for both approaches is that all the selected features are important for at least one of the five classes.

Finally, the z values were computed for the third approach, which considers the classification problem to be binary. Therefore, the z statistic values were obtained directly without employing the OVA structure which is only necessary in multi-class problems. The resulting optimal number of features was chosen according to the mathematical criterion shown in Eqs. 3 and 4, being a total of four. Table 4 shows the selected features for the third approach and the correspondence between these features and the spectral indexes or reflectance values.

Classifier Performance Evaluation

The MLP classifier, trained with the ELM algorithm, was evaluated using the selected features for each approach to the problem on the test set of labelled data. Table 5 shows the results for the first approach using the set of six input features provided by the proposed feature selection methodology. An average success rate of 87.5% is achieved with this approach, this parameter being calculated as the sum of the elements on the main

Table 5 Confusion matrix of the classification of pixels for approach I

Prediction/reality	GS (%)	OS (%)	SC (%)	PD (%)	PI (%)
Green skin	96.91	0.13	0.06	0.03	0.00
Orange skin	2.42	94.69	0.01	2.77	1.22
Scars	0.11	0.03	97.34	0.17	0.83
P. digitatum	0.33	3.02	1.49	74.43	23.93
P. italicum	0.23	2.14	1.10	22.60	74.02

Average success rate=87.5%

Table 6 Confusion matrix of the classification of pixels for approach II

Prediction/reality	GS (%)	OS (%)	SC (%)	PD (%)	PI (%)
Green skin	96.61	0.09	0.03	0.06	0.00
Orange skin	2.40	94.56	0.26	2.72	1.79
Scars	0.08	0.05	98.08	0.00	1.10
P. digitatum	0.82	2.38	0.17	75.90	16.83
P. italicum	0.09	2.92	1.47	21.32	80.28

Average success rate=89.1%

diagonal of the obtained confusion matrix divided by the number of classes.

For the second approach, the evaluation of pixel classification using the set of seven optimal features leads to the confusion matrix shown in Table 7. This approach yields an average success rate of 89.1%.

When comparing the two confusion matrixes (Tables 5 and 6), it can be observed that the number of well-classified pixels of decay classes (PD and PI) for the second approach is greater than that obtained for the first approach. This is due to the fact that these two classes were given the highest weight when obtaining the global relevance for the second approach. Moreover, in the second approach, the classification of pixels with scars (SC) is improved, although to a lesser extent than the classification of the PD and PI classes. It can also be observed that the results for the classification of the sound classes (GS and OS) hardly vary between the two approaches, since these classes are considered of low importance when obtaining the global relevance in the second approach.

Tables 5 and 6 show, in both cases, that the most difficult task in the pixel classification problem is to discriminate the PD class from the PI class, due to the similarity of the damage caused by the two fungi. On the other hand, the low percentage of sound pixels (GS and OS) classified as rotten pixels (PD and PI) in both approaches should also be highlighted. In practice, this is of great importance since most confusion is done between classes that could be grouped into the same category or commercial importance such as decay (PD and PI) and sound (GS and OS).

To conclude the comparison between approaches I and II, from the results obtained, it can be said that the second

Table 7 Confusion matrix of the classification of pixels for approach III

Prediction/reality	Decay (%)	Not decay (%)
Decay	96.48	5.13
Not decay	3.52	94.87

Average success rate=95.5%

approach generally provides better results than the first one, with an increase in the average success rate from 87.5% to 89.1%. This improvement is obtained by taking into account classes with different degrees of importance in the classification problem.

Similarly, Table 7 shows the results of the evaluation of classifier performance for approach III using the set of four input features selected with the proposed method, where an average success rate of 95.5% was achieved. Better results are obtained for this approach, since similar classes are grouped into a single class, thus avoiding the confusion that occurs in the classification of these similar classes.

Conclusions

In this work, a feature selection methodology has been proposed that expands the use of the ROC curve to multi-class classification problems, in order to select a reduced set of features that are effective in the detection of decay in citrus fruits using hyperspectral images. Once the optimal features have been selected, pixels from the images were classified using an MLP trained with a fast new learning algorithm (ELM).

This selection methodology was applied specifically to the detection of decay in citrus fruits caused by two different fungi, *P. digitatum* and *P. italicum*, and other common types of damage, such as scars. The conclusions drawn after performing several tests can be summarised as follows:

- A reduced number of features have been obtained for each of the three approaches to the problem, these numbers being six for the first approach, seven for the second approach and four for the third one. In addition, all the selected features for the first and second approaches are important for at least one of the five classes defined (two kinds of sound skin, two kinds of decay and scars).
- The set of features selected with the second approach provides better classification results than those obtained with the first one and increases the average success rate from 87.5% to 89.1% by taking into account classes with different degrees of importance in the classification problem. On the other hand, as expected, better results were obtained for the third approach (95.5%), specifically aimed at the detection of decay.

References

Aleixos, N., Blasco, J., Navarrón, F., & Moltó, E. (2002). Multispectral inspection of citrus in real time using machine vision and digital signal processors. *Computers and Electronics in Agriculture, 33*(2), 121–137.

Ariana, D. P., Guyer, D. E., & Shrestha, B. (2006). Integrating multispectral reflectance and fluorescence imaging for defect detection on apples. *Computers and Electronics in Agriculture, 50*, 148–161.

Balasundaram, D., Burks, T. F., Bulanona, D. M., Schubert, T., & Lee, W. S. (2009). Spectral reflectance characteristics of citrus canker and other peel conditions of grapefruit. *Postharvest Biology and Technology, 51*, 220–226.

Bennedsen, B. S., Peterson, D. L., & Tabb, A. (2007). Identifying apple surface defects using principal components analysis and artificial neural networks. *Transactions of the ASABE, 50*(6), 2257–2265.

Blanc, P. G. R., Blasco, J., Moltó, E., Gómez-Sanchis, J., Cubero. S. (2009). System for the automatic selective separation of rotten citrus fruit. European patent EP2133157A1.

Blanc, P. G. R., Blasco, J., Moltó, E., Gómez-Sanchis, J., Cubero, S. (2010). System for the automatic selective separation of rotten citrus fruit. United States patent US2010/0121484A1.

Blasco, J., Aleixos, N., Gómez, J., & Moltó, E. (2007). Citrus sorting by identification of the most common defects using multispectral computer vision. *Journal of Food Engineering, 83*(3), 384–393.

Blasco, J., Aleixos, N., Gómez-Sanchis, J., & Moltó, E. (2009). Recognition and classification of external skin damage in citrus fruits using multispectral data and morphological features. *Biosystems Engineering, 103*, 137–145.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition, 30*(7), 1145–1159.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology, 4*(4), 487–504.

Du, C.-J., & Sun, D.-W. (2009). Retrospective shading correlation of confocal laser scanning microscopy beef images for three-dimensional visualization. *Food and Bioprocess Technology, 2*, 167–176.

Eckert, J., & Eaks, I. (1989). *Postharvest disorders and diseases of citrus. The citrus industry.* Berkeley: University California Press.

Farrera-Rebollo, R. R., Salgado-Cruz, M. P., Chanona-Pérez, J., Gutiérrez-López, G. F., Alamilla-Beltrán, L., & Calderón-Domínguez, G. (2011). Evaluation of image analysis tools for characterization of sweet bread crumb structure. *Food and Bioprocess Technology.* doi:10.1007/s11947-011-0513-y.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874.

Gaffney, J. J. (1973). Reflectance properties of citrus fruits. *Transactions of the ASAE, 16*(2), 310–314.

Gitelson, A., Merzyak, M. N., & Lichtenthaler, H. K. (1996). Detection of red-edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology, 148*, 501–508.

Gómez-Sanchis, J., Gómez-Chova, L., Aleixos, N., Camps-Valls, G., Montesinos-Herrero, C., Moltó, E., & Blasco, J. (2008). Hyperspectral system for early detection of rottenness caused by *Penicillium digitatum* in mandarins. *Journal of Food Engineering, 89*(1), 80–86.

Gómez-Sanchis, J., Moltó, E., Camps-Valls, G., Gómez-Chova, L., Aleixos, N., & Blasco, J. (2008). Automatic correction of the effects of the light source on spherical objects. An application to

the analysis of hyperspectral images of citrus fruits. *Journal of Food Engineering, 85*(2), 191–200.

Gómez-Sanchis, J., Martín-Guerrero, J. D., Soria-Olivas, E., Martínez-Sober, M., Magdalena-Benedito, R., & Blasco, J. (2012). Detecting rottenness caused by *Penicillium* in citrus fruits using machine learning techniques. *Expert Systems with Applications, 39*(1), 780–785.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment, 81*, 416–426.

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*, 489–501.

Huang, Y., Kangas, L. J., & Rasco, B. A. (2007). Applications of artificial neural networks (ANNs) in food science. *Critical Reviews in Food Science and Nutrition, 47(2),* 113–126.

Jiménez-Cuesta, M., Cuquerella, J., & Martínez-Jávega, J. M. (1981). Determination of a color index for citrus fruit degreening. In: Proceedings of the International Society of Citriculture, 2, 750–753.

Karimi, Y., Maftoonazad, N., Ramaswamy, H. S., Prasher, S. O., & Marcotte, M. (2009). Application of hyperspectral technique for color classification avocados subjected to different treatments. *Food and Bioprocess Technology.* doi:10.1007/s11947-009-0292-x.

Kim, D. G., Burks, T. F., Qin, J., & Bulanon, D. M. (2009). Classification of grapefruit peel diseases using color texture feature analysis. *International Journal of Agricultural and Biological Engineering, 2*(3), 41–50.

Kondo, N., Ahmad, U., Monta, M., & Murase, H. (2000). Machine vision based quality evaluation of Iyokan orange fruit using neural networks. *Computers and Electronics in Agriculture, 29*, 135–147.

Kurita, M., Kondo, N., Shimizu, H., Ling, P., Falzea, P. D., Shiigi, T., Ninomiya, K., Nishizu, T., & Yamamoto, K. (2009). A double image acquisition system with visible and UV LEDs for citrus fruit. *Journal of Robotics and Mechatronics, 21*(4), 533–540.

Li, J., Rao, X., & Ying, Y. (2011). Detection of common defects on oranges using hyperspectral reflectance imaging. *Computers and Electronics in Agriculture, 78*(1), 38–48.

López-García, F., Andreu-García, A., Blasco, J., Aleixos, N., & Valiente, J. M. (2010). Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach. *Computers and Electronics in Agriculture, 71*, 189–197.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., & Blasco, J. (2011). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology.* doi:10.1007/s11947-011-0725-1.

Magwaza, L. S., Opara, U. L., Nieuwoudt, H., Cronje, P. J. R., Saeys, W., & Nicolaï, B. (2011). NIR spectroscopy applications for internal and external quality analysis of citrus fruit—a review. *Food and Bioprocess Technology..* doi:10.1007/s11947-011-0697-1.

Manickavasagan, A., Jayas, D. S., White, N. D. G., & Paliwal, J. (2010). Wheat class identification using thermal imaging. *Food and Bioprocess Technology, 3*(3), 450–460.

Naidu, R. A., Perry, E. M., Pierce, F. J., & Mekuria, T. (2009). The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two redberried wine grape cultivars. *Computers and Electronics in Agriculture, 66*, 38–45.

Obagwu, J., & Korsten, L. (2003). Integrated control of citrus green and blue molds using *Bacillus subtilis* in combination with sodium bicarbonate or hot water. *Postharvest Biology and Technology, 28*(1), 187–194.

Obenland, D., Margosan, D., Collins, S., Sievert, J., Fjeld, K., Arpaia, M. L., Thompson, J., & Slaughter, D. (2009). Peel fluorescence as a means to identify freeze-damaged navel oranges. *HortTechnology, 19*(2), 379–384.

Palou, L., Smilanik, J., Usall, J., & Viñas, I. (2001). Control postharvest blue and green molds of oranges by hot water, sodium carbonate, and sodium bicarbonate. *Plant Disease, 85*, 371–376.

Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J. C., & Trianni, G. (2009). Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment, 113*(1), S110–S122.

Prechelt, L. (1996). A quantitative study of experimental evaluations of neural network learning algorithms: Current research practice. *Neural Networks, 9*(3), 457–462.

Qin, J., Burks, T. F., Ritenour, M. A., & Bonn, W. G. (2009). Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. *Journal of Food Engineering, 93*, 183–191.

Quevedo, R., & Aguilera. (2010). Color computer vision and stereoscopy for estimating firmness in the salmon (*Salmon salar*) fillets. *Food and Bioprocess Technology, 3*(4), 561–567.

Quevedo, R., Aguilera, J. M., & Pedreschi, F. (2010). Color of salmon fillets by computer vision and sensory panel. *Food and Bioprocess Technology, 3*(5), 637–643.

Rao, C. R., & Mitra, S. K. (1972). *Generalized inverse of matrices and its applications.* New York: Wiley.

Rifkin, R., & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of Machine Learning Research, 5*, 101–141.

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment, 55*, 95–107.

Serrano AJ, Soria E, Martín JD, Magdalena R & Gómez J (2010) Feature selection using ROC curves on classification problems. In: *International Joint Conference on Neural Networks, IJCNN 2010, 28th–30th July 2010.* Barcelona, Spain. Proceedings, pp 1980–1985.

Shih, F. Y. (2010). *Image processing and pattern recognition: Fundamentals and techniques.* New York: Wiley-IEEE.

Slaughter, D. C., Obenland, D. M., Thompson, J. F., Arpaia, M. L., & Margosan, D. A. (2008). Non-destructive freeze damage detection in oranges using machine vision and ultraviolet fluorescence. *Postharvest Biology and Technology, 48*, 341–346.

Sun, D.-W. (Ed.). (2010). *Hyperspectral imaging for food quality analysis and control.* London: Academic.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8(2), 127–150.

Unay, D., & Gosselin, B. (2006). Automatic defect segmentation of 'Jonagold' apples on multi-spectral images: A comparative study. *Postharvest Biology and Technology, 42*, 271–279.

Xu, H. R., Ying, Y. B., Fu, X. P., & Zhu, S. P. (2007). Near-infrared spectroscopy in detecting leaf miner damage on tomato leaf. *Biosystems Engineering, 96*(4), 447–454.

Yang, C. M., Cheng, C. H., & Chen, R. K. (2007). Changes in spectral characteristics of rice canopy infested with brown planthopper and leaffolder. *Crop Science, 47*, 329–335.