Nondestructive internal quality inspection of pear fruit by X-ray CT using machine learning

Tim Van De Looverbosch[1], Md. Hafizur Rahman Bhuiyan[4], Pieter Verboven[1], Manuel Dierick[2], Denis Van Loo[2], Jan De Beenbouwer[3], Jan Sijbers[3], Bart Nicolaï[1]

[1] Division of Mechatronics, Biostatistics and Sensors (MeBioS), Biosystems Department, KU Leuven, Belgium
[2] Tescan XRE nv, Belgium
[3] imec Vision Lab, Department of Physics, University of Antwerp, Belgium
[4] Department of Food Technology and Rural Industries, Bangladesh Agricultural University, Bangladesh
tim.vandelooverbosch@kuleuven.be
Willem De Croylaan 42
3001 Leuven
Belgium

ABSTRACT

To preserve the quality of fresh pear fruit after harvest and deliver quality fruit year-round a controlled supply chain and long-term storage are applied. During storage, however, internal disorders can develop due to suboptimal storage conditions that may not cause externally visible symptoms. This makes them impossible to be detected by current commercial quality grading systems in a reliable and non-destructive way. A combination of a Support Vector Machine coupled with a feature extraction algorithm and X-ray Computed Tomography is proposed to successfully detect internal disorders in 'Conference' and 'Cepuna' pear fruit nondestructively. Classifiers were able to distinguish defective from sound fruit with classification accuracies ranging between 90.2 and 95.1 % depending on the cultivar and number of used features. Moreover, low false positive and negative rates were obtained, respectively ranging between 0.0 and 6.7 %, and 5.7 and 13.3 %. Classifiers trained on 'Conference' data were transferred effectively to the 'Cepuna' cultivar, suggesting generalizability to other cultivars as well. With continuing developments in both hardware and software to increase inspection speed and reduce equipment costs, the method can be implemented in industrial applications, *e.g.,* inline translational X-ray CT.

Keywords:

Support Vector Machine; 3D imaging; image processing; postharvest technology; disorder detection; food grading

1 INTRODUCTION

A controlled supply chain and long-term storage is necessary to deliver quality fruit year-round. However, suboptimal storage conditions can cause severe quality loss by chilling injury, accelerated ripening and senescence, fermentation, stimulated pathogen growth or other physiological decay. The disorders that develop during growth or storage may not cause externally visible symptoms. Examples

are internal browning, watercore, bitter pit or cavities in the fruit tissue (Lu and Lu, 2017; Mercier et al., 2017).

Most commercial quality grading systems are based on external quality attributes only. Sensors for inline inspection of these attributes, most often based on visible and near-infrared (Vis-NIR) spectroscopy, are commercially available, but they do not allow to inspect internal quality attributes, such as absence of internal disorders, in a reliable way (Huang et al., 2020; Nicolaï et al., 2014). In research, however, Vis-NIR has been used successfully to detect internal disorders. Han et al. (2006) reported false positive and negative rates of respectively 4.3 and 5.3 % for brown core detection in 'Yali' pears using transmission vis-NIR spectroscopy at three different locations per sample with precisely aligned pears. Khatiwada et al. (2016) used transmission Vis-NIR spectroscopy to detect internal browning in 'Pink Lady™' apples with four spectra acquired of each fruit and reported a classification accuracy of more than 95 %. Huang et al. (2020) investigated transmission Vis-NIR spectroscopy to detect internal defects in 'Honeycrisp' apples by acquiring spectra from six different locations per fruit. Using the mean spectrum of the six measurements, a classification accuracy of up to 93.1 % was reported. For fruit with a defective tissue area less than 40 % of the total fruit tissue area being rated, the accuracy dropped to 77.3 %. The main challenges for using Vis-NIR for the inspection of the internal quality is that internal defects are mostly not uniformly distributed in the fruit. This causes the need for long exposure times, the requirement for a fixed orientation, or the need for multiple measurements from different positions (Lu and Lu, 2017). Additionally, the measured spectrum is hard to interpret and, moreover, depends on the fruit size, cultivar and season. Therefore, large amounts of data are required to perform calibration (Bobelyn et al., 2010; Nicolaï et al., 2007).

Several visualization techniques have been proposed as an alternative to assess internal quality and disorders non-destructively. These include magnetic resonance imaging (MRI), X-ray radiography and X-ray computed tomography (CT) (Arendse et al., 2018; Kotwaliwale et al., 2014; Nicolaï et al., 2014; Srivastava et al., 2018; Wang et al., 2018). For MRI, the remaining concerns are low image acquisition speed due to physical constraints, the need for a sufficiently powerful and homogenous magnetic field, high equipment costs, electromagnetic inference and motion artefacts (Colnago et al., 2014; Srivastava et al., 2018). X-ray radiography provides 2D transmission images of the scanned fruit with a contrast that scales with the accumulated amount of absorbed X-rays by the product. It can be easily implemented inline in the form of an X-ray line-scanner (Casasent et al., 1998; Jiang et al., 2008; Karunakaran et al., 2004; Kim and Schatzki, 2001, 2000; Kotwaliwale et al., 2007; Shahin et al., 2001; van Dael et al., 2016). X-ray radiography image contrast may, however, suffer from effects of fruit shape, volume and internal structure such that internal defects can be less prominent in the image

when they are, e.g., shadowed by the core of the fruit. Good detection and classification are then jeopardized unless advanced approaches are pursued (van Dael et al., 2018, 2017, 2015).

X-ray CT delivers 3D images of the X-ray absorption contrast, thereby resolving shape and volume of the fruit while providing better spatial information on the internal fruit structure and disorders than in radiographies. CT has been successfully used in research to analyze the internal quality of fruit and vegetables (Donis-González et al., 2016a, 2016b, 2014; Herremans et al., 2014, 2013; Jarolmasjed et al., 2016; Lammertyn et al., 2003a, 2003b; Muziri et al., 2016; Si and Sankaran, 2016). Yet, it is still challenging to use CT efficiently for high-throughput quality inspection, which limits its usage in industrial applications to the inspection of high value products at a relatively slow rate, or to its usage in research and development stages (Buratti et al., 2018; Wevers et al., 2018). However, recent developments in both reconstruction algorithms and hardware have made it possible to perform X-ray CT measurements inline by translating and rotating the sample at the same time, i.e., translational X-ray CT. To facilitate fast inline CT scanning new hardware equipment must be developed. However, these are technical challenges that, with enough resources available, can already be overcome. Moreover, by using prior knowledge about the products to train dedicated reconstruction algorithms, reconstructions can be made with far less projections than normally required (Alves Pereira et al., 2017; De Schryver et al., 2016; Janssens et al., 2018, 2016). The concerns that CT shares with MRI relating inspection speed and inline application might therefore be partially resolved.

To respond to these promising advancements, a nondestructive method for automated internal quality grading of pear fruit using machine learning is proposed, aimed to be implemented in inline translational X-ray CT. Hereto, a binary linear Support Vector Machine (SVM) is trained and tested on labeled X-ray CT reconstructions of control and defective pear fruit of the cultivars *Pyrus communis* L. cv. 'Conference' and *Pyrus communis* L. cv. 'Cepuna'. 'Conference' is one of the most important commercial cultivars in Europe, represents almost 90 % of the acreage of Belgian pears (Statbel, 2018), and is known to be susceptible to internal browning (Franck et al., 2007). 'Cepuna' is a cross between 'Conference' and 'Doyenné d'Hiver' and used for testing the transferability of the method to other cultivars.

2 MATERIALS AND METHODS

2.1 Pear fruit and long-term storage

'Conference' and 'Cepuna' pears were respectively harvested on 14 and 25 September 2017 and delivered by a growers member of the Flemish fruit cooperatives BFV and Belorta (Belgium), respectively. Starting from the harvest date, the fruit was stored for six months following two treatments, with approximately 50 kg fruit per treatment. In the first treatment, the storage conditions

were set according to the recommendations of the Flanders Centre of Postharvest Technology (VCBT, Leuven, Belgium) for commercial sale (Ultra Low Oxygen treatment, ULO) to deliver control fruit without internal disorders (VCBT, 2017). Herein, the temperature, O_2 and CO_2 partial pressures were set to -1.0 °C, 3.0 kPa and 0.7 kPa, respectively. Prior to the ULO storage, fruit following this treatment underwent an acclimatization period of 21 d at -1.0 °C. In the second treatment suboptimal storage conditions, based on the findings of (Lammertyn et al., 2000), were applied to deliver fruit with internal disorders. Herein, the temperature, O_2 and CO_2 partial pressures were respectively -1.0 °C, 1.0 kPa and 5.0 kPa. A low O_2 partial pressure causes hypoxia in the fruit. In combination with increased CO_2 partial pressure, this promotes the shift from respiration to fermentation, resulting in a limited availability in energy and an imbalance between oxidative and reductive processes. As such, cell membranes are degraded by reactive oxygen species, leading to cell leakage and cell death, which manifest as internal browning and cavity formation (Franck et al., 2007; Pedreschi et al., 2009; Veltman et al., 2003).

2.2 X-ray CT scans and data labeling

After approximately 6 months, the fruit was removed from storage on 2018-02-27 at the end of the day. Fruit were acclimatized to room conditions before X-ray CT scanning the next day. Minimally 50 fruit per treatment were randomly selected and scanned individually. The fruit was scanned with their stalk-calyx axis approximately aligned with the rotation axis of the scanner. To stabilize the samples during scanning, the fruit was placed on a sample holder consisting of three styrofoam cones glued on a stainless-steel plate which was mounted on top of the rotation table. The system comprised a micro-focus L9181 X-ray source (Hamamatsu Photonics, Hamamatsu, Japan) and a 1512 Dexela CMOS Flat Panel X-ray Detector (PerkinElmer, Waltham, Massachusetts, USA). The rotation table and detector were placed at respectively 674.8 mm and 784.2 mm from the source. The X-ray projections did not fit entirely in the X-ray detector frame. Therefore, two scans per fruit were performed at different heights and stacked together to reconstruct the whole fruit in the CT volume. The scans were performed with a source voltage of 130 kV at 300 mA and pixel size of 598.4 µm. The exposure time was 80 ms. An aluminum filter of 1 mm thickness was used to improve the contrast in the radiographic projections that were obtained with an angular step of 0.9° and were 242 × 192 pixels in size. The samples were rotated over 360° around the central rotation axis of the scanner, resulting in 400 projections. For the acquisition, ACQUILA software was used (Tescan XRE nv, Ghent, Belgium). A 3D image was reconstructed of each fruit with the filtered back-projection algorithm using the ACQUILA-RECON reconstruction software (Tescan XRE nv, Ghent, Belgium). The resulting tomographs had a size of 241 × 241 × 309 voxels, with each isotropic voxel measuring 514.9 × 514.9 × 514.9 μm^3. In total, scanning (32 s/scan), moving the sample stage down and starting the second scan (2 s), stacking and

reconstruction (23 s) amounted on average to 1 min and 30 s per sample. The samples were assigned a ground truth label (*'healthy'* or *'defective'*) by visual inspection of the CT reconstruction of each fruit.

Figure 1 shows the experimental X-ray CT setup and a cut-open image and orthogonal slices through the CT volume of a *'Cepuna'* pear severely affected by internal browning. Internal browning can be observed in the lower intensity regions on the CT slices.

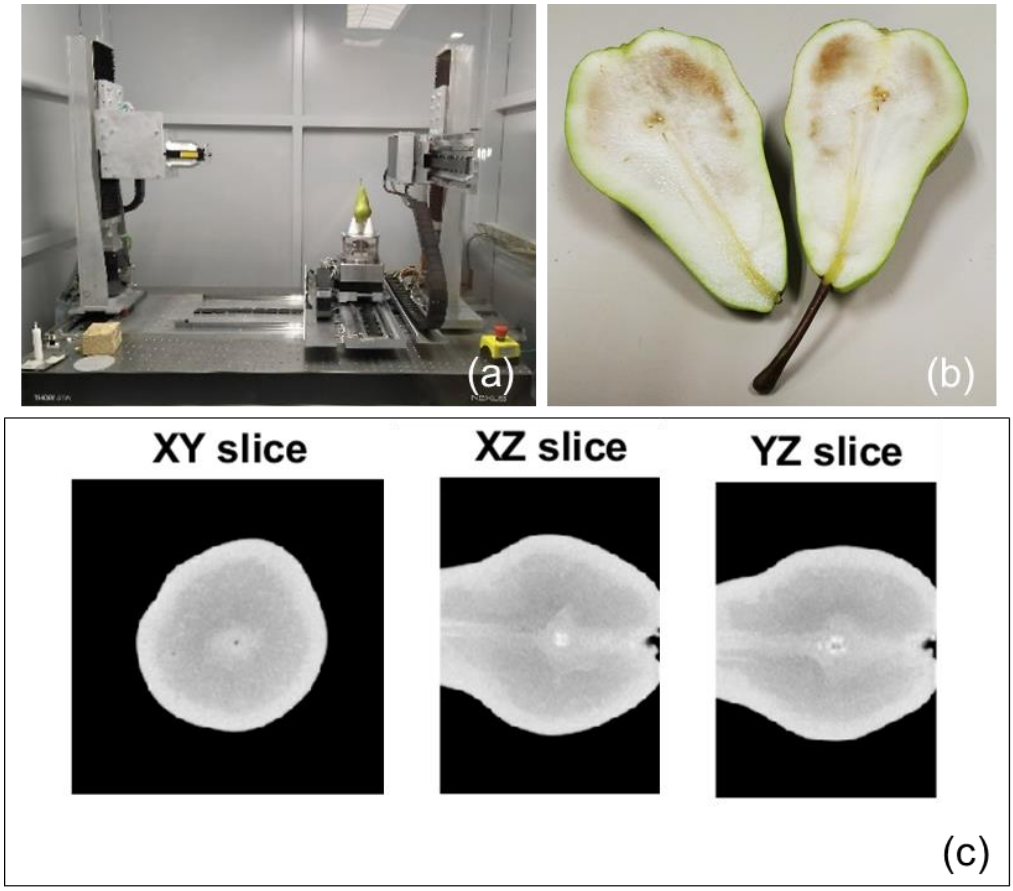

Figure 1: Experimental setup and a 'Cepuna' pear's cut-open image and CT scan. (a) Experimental setup with X-ray source (left), and mobile X-ray detector (right) and rotation stage (middle); (b) image of a cut-open 'Cepuna' pear affected by internal browning; (c) orthogonal slices through the CT volume of the same fruit. The XZ and YZ slices are zoomed in on the region affected by internal browning.

2.3 Internal disorder detection method on CT images

An algorithm was developed to perform internal disorder detection for pear fruit using the CT images. Recent developments in the medical imaging field reported interesting results for disease detection in CT and MRI data using deep learning based segmentation methods (Lee et al., 2017; Shen et al., 2017). However, these approaches typically require many manually labelled samples for training. Due to limited number of samples and the high cost of manually labelling them, a more classical machine learning approach was chosen. First, a feature extraction algorithm was developed to get valuable quantities, or features, from the 3D image datasets (see section 2.3.1). Subsequently, the features were statistically compared between the cultivars and classes (see section 2.3.2). Thereafter, support

vector machines (SVM) were trained separately on the 'Conference' feature dataset to classify the fruit. Then, it was investigated whether features could be eliminated while minimizing the reduction in the classification performance. Finally, to test the generalizability of the method, the classifier trained on the 'Conference' data was validated on fruit of the 'Cepuna' cultivar and compared with classifiers trained on the combined dataset (see section 2.3.3). All code was written with MATLAB using the Image Processing and Statistics and Machine Learning Toolboxes (MATLAB, 2019a).

2.3.1 Feature extraction algorithm

A feature extraction algorithm was developed to extract 10 features from the CT volume of pear fruit and produced 2 feature datasets, one for each cultivar. To extract the features, five 3D binary masks were generated indicating different regions of the fruit (see Figure 2). Each 3D binary mask specified whether a certain voxel belonged to a certain segment of the volume (value = 1) or not (value = 0). First, a binary mask that indicated which voxels were part of the fruit tissue (tissue mask) was generated using a 3D global Otsu-threshold (Otsu, 1979). Second, a fruit mask was generated by filling up all internal holes of the tissue mask so that voxels outside and inside the fruit had the values 0 and 1, respectively. Third, an internal air mask, only including voxels part of holes, was obtained by subtracting the tissue mask from the fruit mask. Finally, a low-density and a high-density tissue mask were generated by applying a 3D adaptive threshold on all tissue voxels based on the local mean intensity in a 31 x 31 x 31 voxel neighborhood. In pome fruit, tissue of higher density can typically be observed around the core and in the surface region. In between those regions, typically a higher porosity can be found (Nugraha et al., 2019). Low-density and high-density tissue regions are thus generally always present, however, a large difference between those regions can indicate the occurrence of water loss due to internal tissue breakdown. In the reconstructed CT volume, voxels with a relative low intensity value had a lower X-ray attenuation, and thus lower density (higher porosity), than voxels with a higher intensity value.

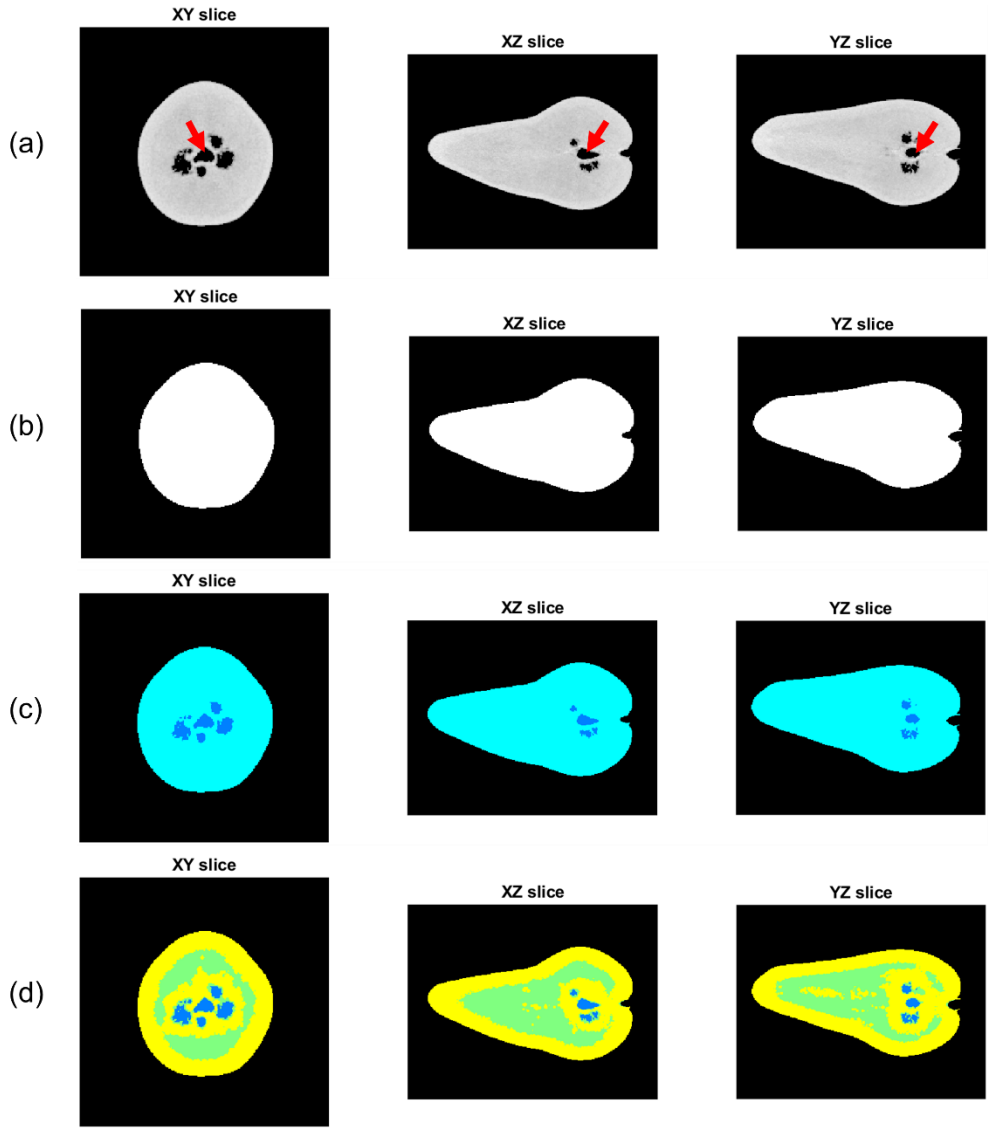

Figure 2: Orthogonal slices through the original grayscale CT reconstruction and generated 3D binary masks of a 'Conference' sample with disorders. (a) Orthogonal slices through the original grayscale CT reconstruction of the sample; cavities can be observed around the core (core indicated by red arrow); (b) orthogonal slices through the fruit mask (white); (c) orthogonal slices through the tissue (cyan) and internal air masks (dark blue); (d) orthogonal slices through the internal air (dark blue), low-density tissue (green) and high-density tissue (yellow) masks. Each 3D binary mask indicates whether a certain voxel belongs to a certain segment of the volume (value = 1) or not (value = 0).

Features were extracted using the generated masks. By subtracting the tissue mask from the fruit mask and counting the number of remaining voxels, the internal air volume could be calculated. As a first feature, the internal air volume normalized for the total fruit volume was used. For the second to ninth feature, the mean and standard deviation of the intensities of fruit voxels, tissue voxels, low density tissue voxels and high-density tissue voxels were calculated by using the fruit, tissue, low density tissue and high-density tissue mask, respectively. As a final feature, the Kolmogorov-Smirnov test statistic (KS-value) of the Two-Sample Kolmogorov-Smirnov Test between the cumulative intensity distributions of the low- and high-density tissue voxels was used (Massey, 1951; MATLAB, 2019b). Here, the KS-value was interpreted as a measure of homogeneity of the fruit tissue by comparing the

intensity distributions of both regions. A lower KS-value indicates that the low-density and high-density tissue regions are of similar density, suggesting that internal tissue breakdown such as browning is less probable (Franck et al., 2007). This is illustrated for a *'defective'* and *'healthy'* 'Conference' fruit in Figure 3. The feature datasets were centered and scaled using the corresponding column mean and standard deviation.

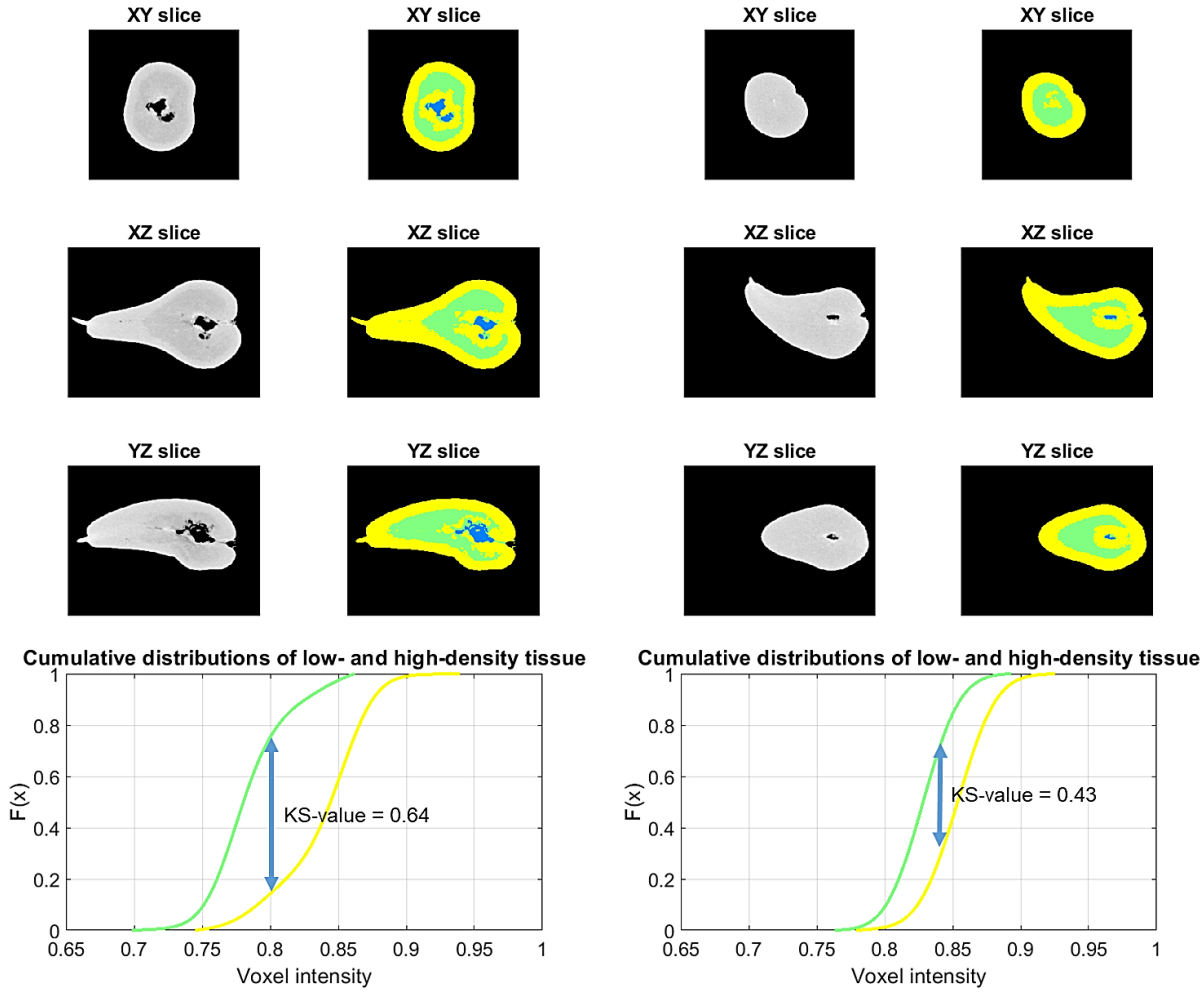

Figure 3: Orthogonal slices through CT volume, cumulative intensity distributions and KS-value of the low- (green) and high-density tissue (yellow) in a 'defective' (left) and 'healthy' *(right) 'Conference' pear. The* 'healthy' *sample has more similar intensity distributions and lower KS-value compared to the* 'defective' *sample*

2.3.2 Statistical feature comparison

A quantitative feature comparison was performed to explore the data, investigate differences between the cultivars or classes and infer relevant features for classification. Hereto, it was tested if the normal distributions of the features were significantly different between the 'Conference' and 'Cepuna' cultivars on the one hand and the *'defective'* and *'healthy'* classes on the other hand using a Two-Sample t-test at the 5% significance level. Moreover, the linear correlation coefficients between all features of the 'Conference' feature dataset were calculated.

2.3.3 Binary linear Support Vector Machine classifiers

Using the 'Conference' feature dataset, a binary linear support vector machine (SVM) was trained and evaluated to classify the fruit. A linear classifier was chosen over non-linear approaches because of its simplicity in terms of the number of parameters, because it allows to interpret the importance of each feature for classification and because it is less prone to overfitting compared to non-linear methods. In contrast to other classifiers, *e.g.,* k-nearest neighbors, the training data is afterwards no longer needed for making predictions.

Confusion matrices were used to present the classification results with true positives (the correctly classified fruit with internal disorders) and true negatives (the correctly classified fruit without internal disorders) shown as a percentage on the matrix diagonal. The false positives and false negatives are shown as a percentage on the bottom left and the top right, respectively.

Thereafter it was investigated whether the number of features used by the classifier could be reduced without losing classification performance. Hereto the SVM recursive feature elimination method (SVM RFE) as described by Guyon et al. (2002) was used in which the importance of each feature relative to the other features was evaluated based on the weights that define the decision boundary of the SVM in feature space. The higher the squared weight value, the more important the corresponding features is for classification. Note that this is only possible when using a linear kernel in the SVM, as for other kernels the data is transformed to another space not related to the original input space. In practice, a series of classifiers was trained and evaluated on the 'Conference' dataset using 5-fold cross-validation (further referred to as the *'Conference' based SVMs*). In each iteration, the feature with the lowest squared weight value was eliminated. By tracking the classification accuracy and false positive and negative rates, a decision was made on which features were the most critical and which classifier should be used.

Next, the generalizability of the trained classifier to other cultivars was evaluated and it was investigated whether the generalizability would increase with a reduction in the number of features. Hereto, the series of trained classifiers was validated on the feature dataset of the 'Cepuna' cultivar. Finally, the 'Conference' and 'Cepuna' datasets were combined and the performance of the series of *'Conference' based SVMs'* was compared with two series of SVMs retrained on this combined dataset. The first series was forced to use the same features as the *'Conference' based SVMs'* in each iteration, while in the second series the feature elimination algorithm decided which features were used.

The runtime for feature extraction and classification was on average 2.3 s per sample on a quad-core 3.8 GHz processor with 32 GB of RAM memory.

3 RESULTS

3.1 X-ray micro CT reconstructions and labeled datasets

For 'Conference', 102 samples were scanned of which 42 and 60 fruit were assigned a *'healthy'* and *'defective'* label, respectively, from expert inspection of the CT images. For 'Cepuna', 15 *'healthy'* and 87 *'defective'* fruit were observed in the 102 scanned samples.

Examples of orthogonal slices and grayscale intensity profiles through CT reconstructed volumes of *'healthy'* and *'defective'* 'Conference' and 'Cepuna' fruit are shown in Figure 4. In the *'healthy'* fruit (Figure 4, rows b and d) a gradient in voxel intensity can be observed from the center to the fruit surface. Higher intensities due to higher tissue density were observed around the core. When moving from the core towards the fruit surface, the intensities first decreased and thereafter increased again closer to the surface, confirming the observed density distributions from other research (Nugraha et al., 2019). The *'defective'* 'Conference' fruit (Figure 4, row a) showed regions of lower voxel intensities that were affected by internal browning (Franck et al., 2007; Lammertyn et al., 2003a; van Dael et al., 2017). Severe internal browning resulted in cavity formation, which was observed around the core and stalk-calyx axis. The *'defective'* 'Cepuna' fruit (Figure 4, row c) were also affected by internal browning, but cavity formation was far less severe. In the grayscale intensity profiles of the *'defective'* fruit, the regions affected by internal browning caused a stronger slope compared to those of the *'healthy'* fruit.

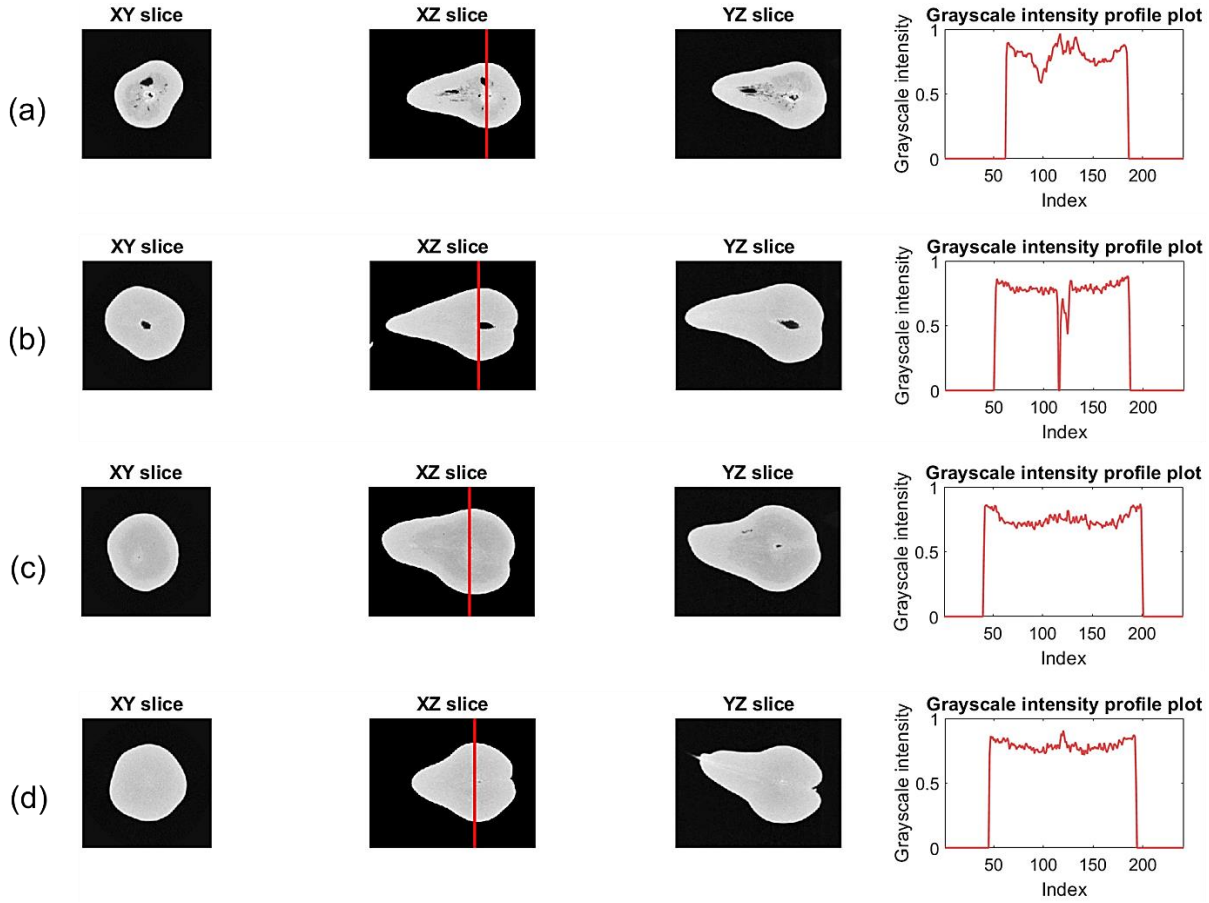

Figure 4: Column 1-3: Orthogonal slices through the CT reconstructions of 'defective' *and 'healthy' 'Conference' and 'Cepuna' fruit. Column 4: Grayscale intensity profile through the widest position of the fruit in the XZ slices. (a) 'defective' 'Conference' pear; (b) 'healthy' 'Conference' pear; (c)* 'defective' *'Cepuna' pear; (d) 'healthy' 'Cepuna' pear.*

3.2 Quantitative feature comparison

The average of the extracted features and their corresponding standard deviations for 'Conference' and 'Cepuna' *'defective'* and 'healthy' fruit are shown in Table 1. Using a Two-Sample t-test at the 5 % significance level, it was tested if the distributions of feature values were significantly different between the 'Conference' and 'Cepuna' cultivars on one hand and the *'defective'* and *'healthy'* classes on the other hand. Figure 5 presents the linear correlation coefficients (R) between all features of the 'Conference' feature dataset.

271 *Table 1: The average of feature values and their corresponding standard deviations for 'Conference' and 'Cepuna' 'defective'*
272 *and 'healthy' fruit. Different letters in superscript indicate significantly different normal distributions at the 5% significance*
273 *level using the Two-Sample t-test.*

Feature	Label	'Conference'	'Cepuna'
Normalized cavity volume [%]	Defective	0.752 ± 0.745^{c}	1.205 ± 1.297^{d}
	Healthy	0.195 ± 0.156^{b}	0.002 ± 0.002^{a}
Mean fruit intensity	Defective	0.829 ± 0.015^{c}	0.802 ± 0.017^{a}
	Healthy	0.835 ± 0.012^{d}	0.824 ± 0.008^{b}
Std fruit intensity	Defective	0.091 ± 0.025^{c}	0.102 ± 0.038^{c}
	Healthy	0.066 ± 0.008^{b}	0.053 ± 0.001^{a}
Mean tissue intensity	Defective	0.835 ± 0.013^{c}	0.812 ± 0.013^{a}
	Healthy	0.836 ± 0.012^{c}	0.824 ± 0.008^{b}
Std tissue intensity	Defective	0.062 ± 0.004^{c}	0.064 ± 0.006^{d}
	Healthy	0.056 ± 0.002^{b}	0.053 ± 0.001^{a}
Mean low-density tissue intensity	Defective	0.814 ± 0.019^{c}	0.782 ± 0.023^{a}
	Healthy	0.825 ± 0.013^{d}	0.806 ± 0.010^{b}
Std low-density tissue intensity	Defective	0.027 ± 0.060^{c}	0.039 ± 0.020^{d}
	Healthy	0.020 ± 0.002^{a}	0.023 ± 0.002^{b}
Mean high-density tissue intensity	Defective	0.856 ± 0.012^{c}	0.838 ± 0.013^{a}
	Healthy	0.854 ± 0.011^{c}	0.846 ± 0.006^{b}
Std high-density tissue intensity	Defective	0.025 ± 0.003^{c}	0.031 ± 0.005^{d}
	Healthy	0.023 ± 0.002^{a}	0.023 ± 0.001^{a}
KS-value	Defective	0.578 ± 0.116^{b}	0.586 ± 0.129^{b}
	Healthy	0.506 ± 0.048^{a}	0.610 ± 0.038^{b}

274

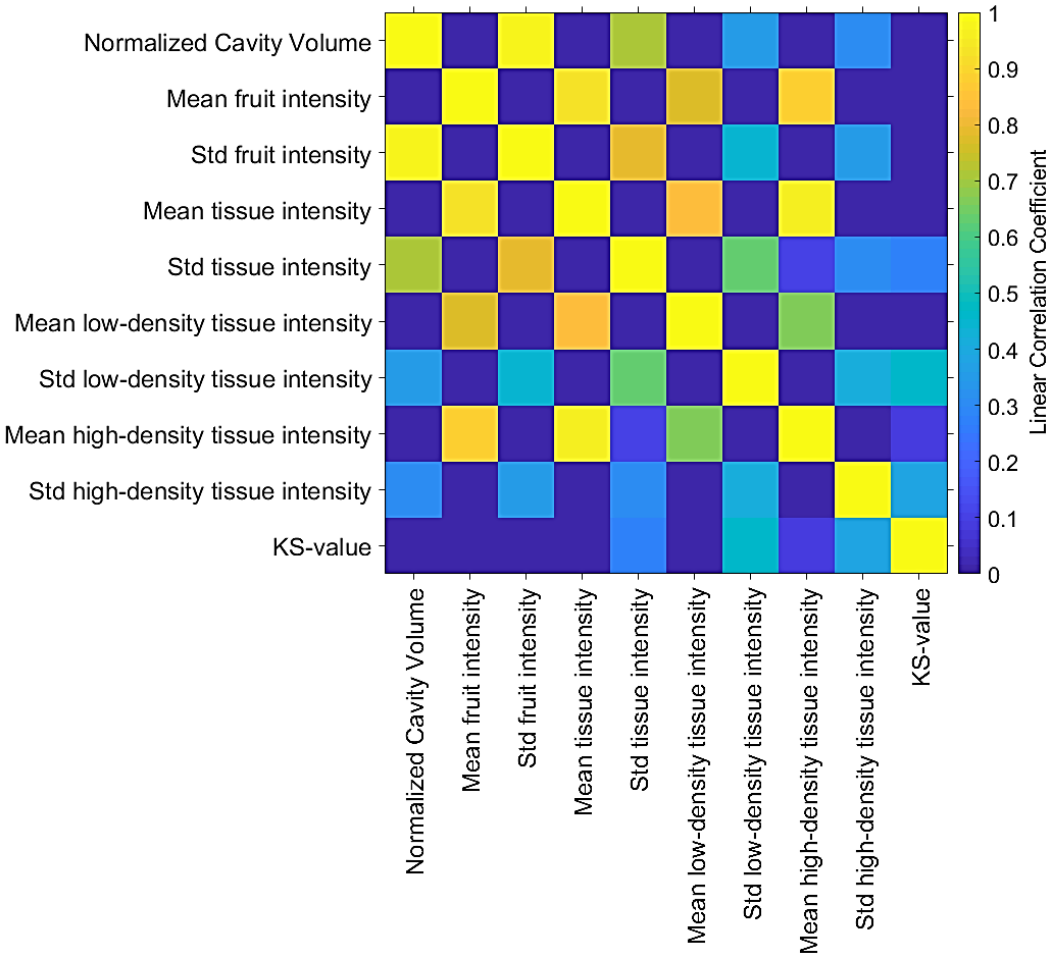

Figure 5: Linear Correlation Coefficients between all features of the 'Conference' dataset.

The features *'Normalized cavity volume'* and *'Std fruit intensity'* were highly correlated (R = 0.98). This was expected as fruit with more or large cavities would also have a high variability in voxel intensity. A lower linear correlation was found between the features *'Normalized cavity volume'* and *'Std tissue intensity'* (R = 0.71). The *'Std tissue intensity'* only considers non-cavity voxels, but fruit with a relatively high number of cavities could have more partial volume artefacts (the loss of contrast in voxels that are occupied by multiple types of tissue due to insufficient resolution), or internal browning and thus a higher variability in fruit tissue intensity.

We also observed that the features *'Mean fruit intensity'*, *'Mean tissue intensity'*, *'Mean low-density tissue intensity'* and *'Mean high-density tissue intensity'* were highly correlated, with linear correlation coefficients ranging between 0.66 and 0.95. Not surprisingly, the correlations were higher between the features of which the regions indicated by the 3D binary masks were more similar, *e.g., 'Mean fruit intensity'* and *'Mean tissue intensity'* had a higher correlation than *'Mean low-density tissue intensity'* and *'Mean high-density tissue intensity'*.

Obviously, as the fruit and tissue masks only differ in the cavity voxels, a rather high correlation (R = 0.78) was found between *'Std fruit intensity'* and *'Std tissue intensity'*. *'Std tissue intensity'* and *'Std low-density tissue intensity'* had a linear correlation of R = 0.64, while *'Std high-density tissue intensity'* and *'KS-value'* where not highly correlated with other features.

3.3 Classification results

3.3.1 'Conference' based SVM

The classifier trained on the whole 'Conference' feature dataset, comprising 60 'defective' and 42 'healthy' samples, reached an average classification accuracy of 92.2 % for 'Conference' in a 5-fold cross validation with an 88.3 % true positive and a 97.6 % true negative rate, respectively. The confusion matrix with classification results for 'Conference' is shown in Table 2.

Table 2: Confusion matrix with average classification results of the 'Conference' specific classifier on 'Conference' in 5-fold cross validation.

		Predicted	
		Defective	Healthy
Ground truth	Defective	88.3 %	11.7 %
	Healthy	2.4 %	97.6 %
	Overall classification accuracy:	92.2 %	

The weights that determine the separating plane are shown in Figure 6. The features with a high absolute value of the weight are the most important for determining the class of a fruit. The top three features were *'Std tissue intensity'*, *'Std high-density tissue intensity'* and *'Std low-density tissue intensity'*. Features that measure variability rather than absolute values had higher absolute weights and, thus, were more important for classifying pear fruit. Moreover, fruit with higher values for these features were more likely to be classified as *'defective'* fruit, *i.e.,* the positive class, due the positive corresponding weights. Both the weights of *'Mean tissue intensity'* and *'KS-value'* features were rather insignificant.

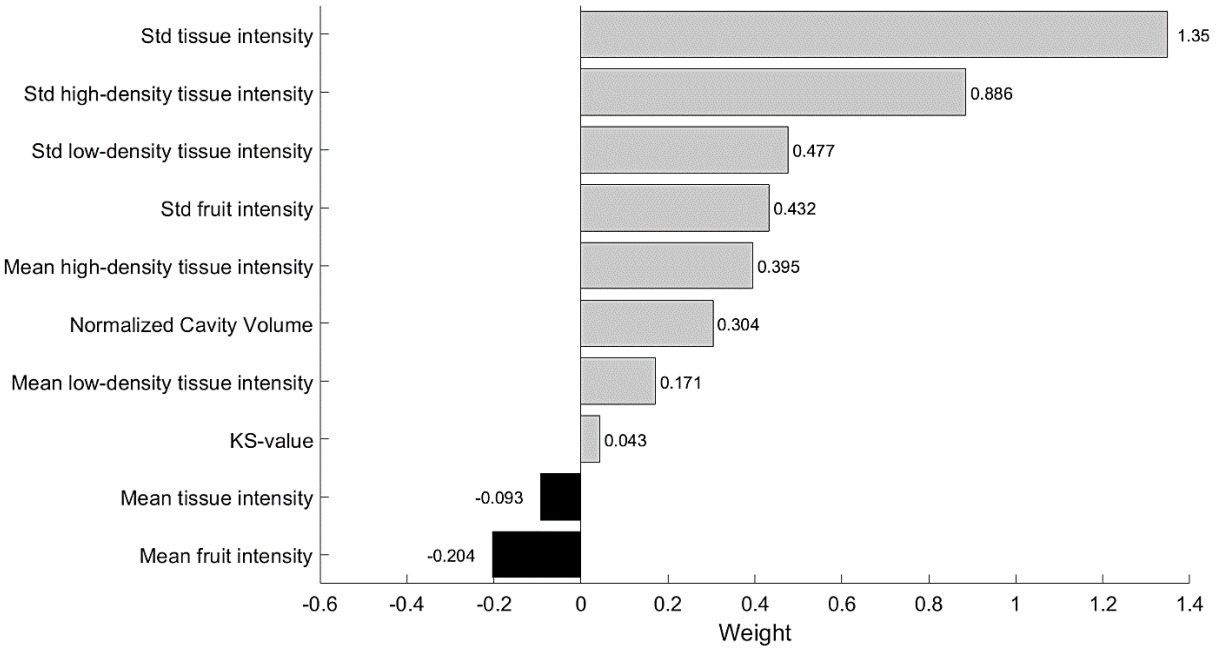

Figure 6: Weights of the 'Conference' based SVM' sorted by descending weight value.

3.3.2 Feature selection

From the previous results, it was observed that not all features were equally important for classification. Some features, *e.g.,* '*KS-Value'*, *'Mean Tissue Intensity'* and *'Mean low-density tissue intensity'*, have relatively low weights compared to others (see Figure 6). As explained in section 2.3.3, the SVM RFE method was applied to select the most relevant features. A series of classifiers was trained and evaluated on the 'Conference' dataset and in each iteration the feature with the lowest squared weight value was removed for the next iteration. The resulting features used by each classifier are shown in Figure 7.

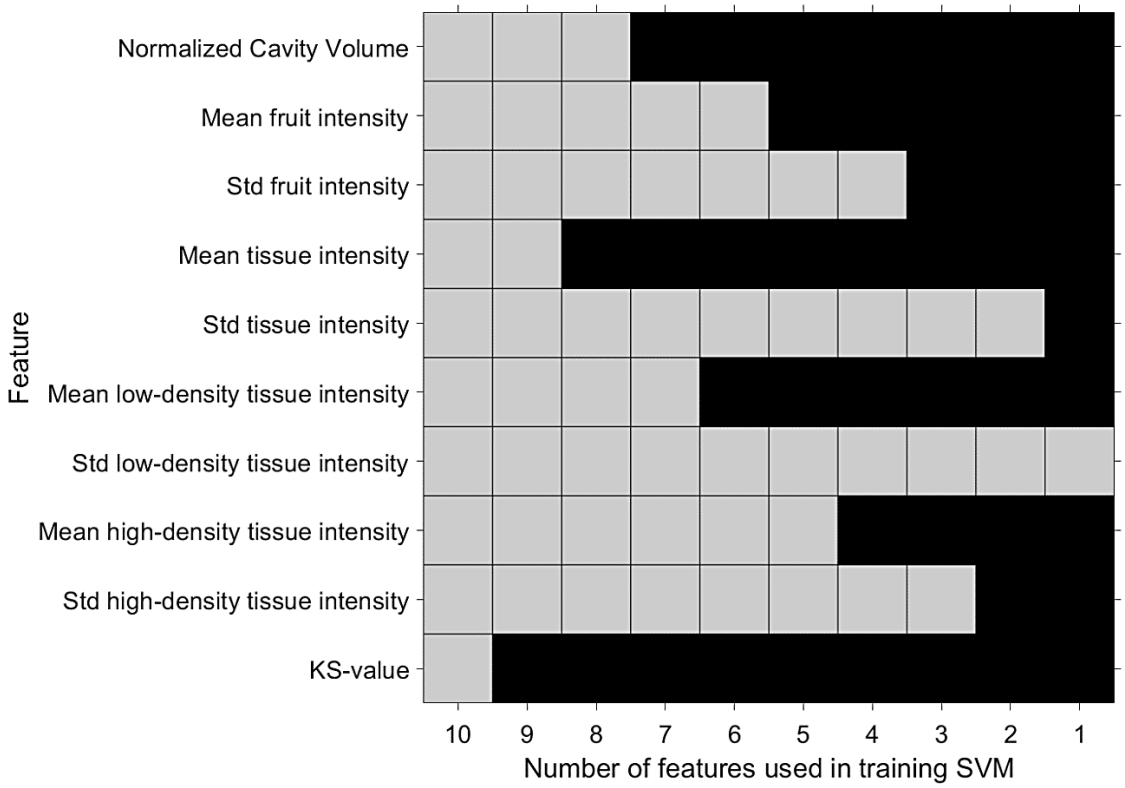

Figure 7: Plot of features used for training each SVM on the 'Conference' dataset. Each column represents a classifier in a series of classifiers, in which the number of features allowed to be used decreased from the left (10 features) to the right (1 feature). Every column thus shows the features used by a classifier in the series, while each row shows in which classifiers a certain feature was used. A grey tile indicates a feature was used, while a black tile indicates a feature was eliminated for the specific SVM in the feature elimination procedure.

Figure 8 shows a plot of the classification accuracy, true positive, true negative, false positive and false negative rate of the SVM series trained with 1 up to 10 features of the 'Conference' dataset in a 5-fold cross-validation.

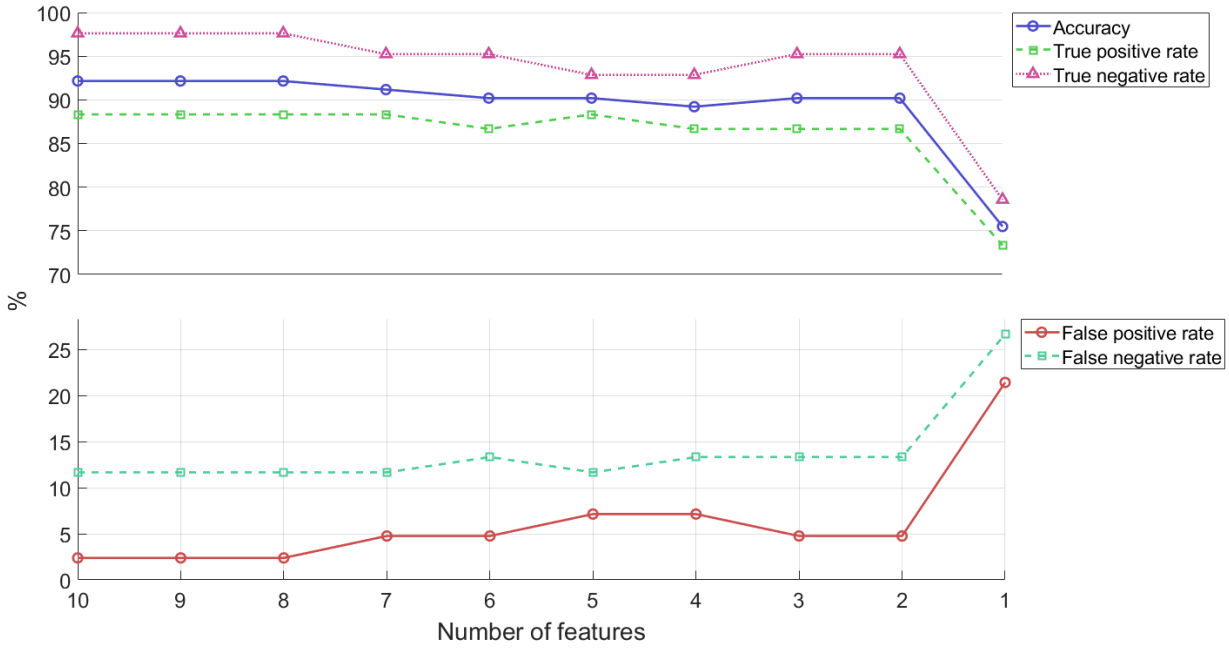

Figure 8: Plot of classification accuracy, true positive rate, true negative rate, false positive rate and false negative rate of SVMs trained with 1 and up to 10 features of the 'Conference' dataset.

From Figure 8 it can be observed that the performance metrics only slightly decreased up until the last SVM trained with only 1 feature, for which the accuracy dropped to 75.5 %. The SVM trained with two features, (*'Std tissue intensity'* and *'Std low-density tissue intensity',* see Figure 7) still achieved an accuracy of 91.2 % and the false positive rate of 4.8 % while the false negative rate increased slightly from the 11.7 % of the best classifier to 13.3 %.

3.3.3 Validating the 'Conference' based SVMs' on the 'Cepuna' cultivar

To test the generalizability to other cultivars, the series of 'Conference' based SVMs' was validated on the 'Cepuna' cultivar without retraining the classifiers on the 'Cepuna' data. The 'Cepuna' dataset comprised 87 *'defective'* and 15 *'healthy'* fruit. The ten-features classifier reached a good classification accuracy of 95.1 % for 'Cepuna' with 94.3 % true positive and 100.0 % true negative rate, respectively. The confusion matrix with classification results for 'Cepuna' is shown in Table 3.

Table 3: Confusion matrix with classification results of the ten-features 'Conference' based SVM' on 'Cepuna' without retraining the SVM.

		Predicted	
		Defective	Healthy
Ground truth	Defective	94.3 %	5.7 %
	Healthy	0.0 %	100.0 %
	Overall classification accuracy:	95.1 %	

When testing the *'Conference' based SVM* series trained with 1 and up to 10 features on the 'Cepuna' dataset, the classification performance remained the same for the classifiers using between ten and five features. For the classifiers using between 4 and 2 features, the false positive rate increased from

0.0 % to 6.7 %, while the accuracy remained 95.1 % and the false negative rate reduced from 5.7 % to 4.7 %. With only one feature, the last classifier had a false positive rate of 80.0 %. However, due to the low number of *'healthy'* samples in the 'Cepuna' dataset (fifteen), the accuracy only dropped to 86.7 %.

3.3.4 Testing the 'Conference' based SVMs' and retrained classifiers on the combined dataset

The series of *'Conference' based SVMs'* was tested on the combined dataset. The confusion matrix with classification results of the ten-features classifier is shown in Table 4.

Table 4: Confusion matrix with average classification results of the ten-features 'Conference' based SVM' on the combined dataset.

		Predicted	
		Defective	Healthy
Ground truth	Defective	91.8 %	8.2 %
	Healthy	1.7 %	98.3 %
	Overall classification accuracy:	93.6 %	

Like the previous results, the classification accuracy, true positive rate, true negative rate, false positive rate and false negative rate of the *'Conference' based SVMs'* classifiers using between ten and two features tested on the combined dataset was similar. The two-feature *'Conference' based SVM'* achieved an accuracy of 93.1 % and a false positive and false negative rate of 3.5 % and 8.2 %, respectively.

Next, a first series of SVMs was retrained on the combined dataset but was forced to use the same features as their corresponding *'Conference' based SVM'* (see Figure 7). The SVMs were thus only allowed to change the weight associated to a certain feature. The classifiers using ten and two features achieved the same classification metric scores, with an accuracy of 92.7 % and false positive and false negative rate of 5.3 and 8.2 %, respectively. However, the accuracy and false positive rate were slightly worse compared to the two-feature *'Conference' based SVM'*, that achieved the same false negative rate with an accuracy of 93.1 % and false positive rate of 3.5 %.

Finally, a second series of SVMs was retrained on the combined dataset which was now allowed change the selected features at each iteration. The used features are shown in Figure 9. The ten-feature classifier reached an accuracy of 92.7 % and false positive and false negative rate of 5.3 and 8.2 %. The two-feature classifier scored an accuracy of 91.2 % with a false positive and false negative rate of 12.3 and 7.5 %, respectively. The latter classifier used the features *'Mean fruit intensity'* and *'Mean high-density tissue intensity'* in contrast to the two-feature *'Conference' based SVM'* that used the features *'Std tissue intensity'* and *'Std low-density tissue intensity'*.

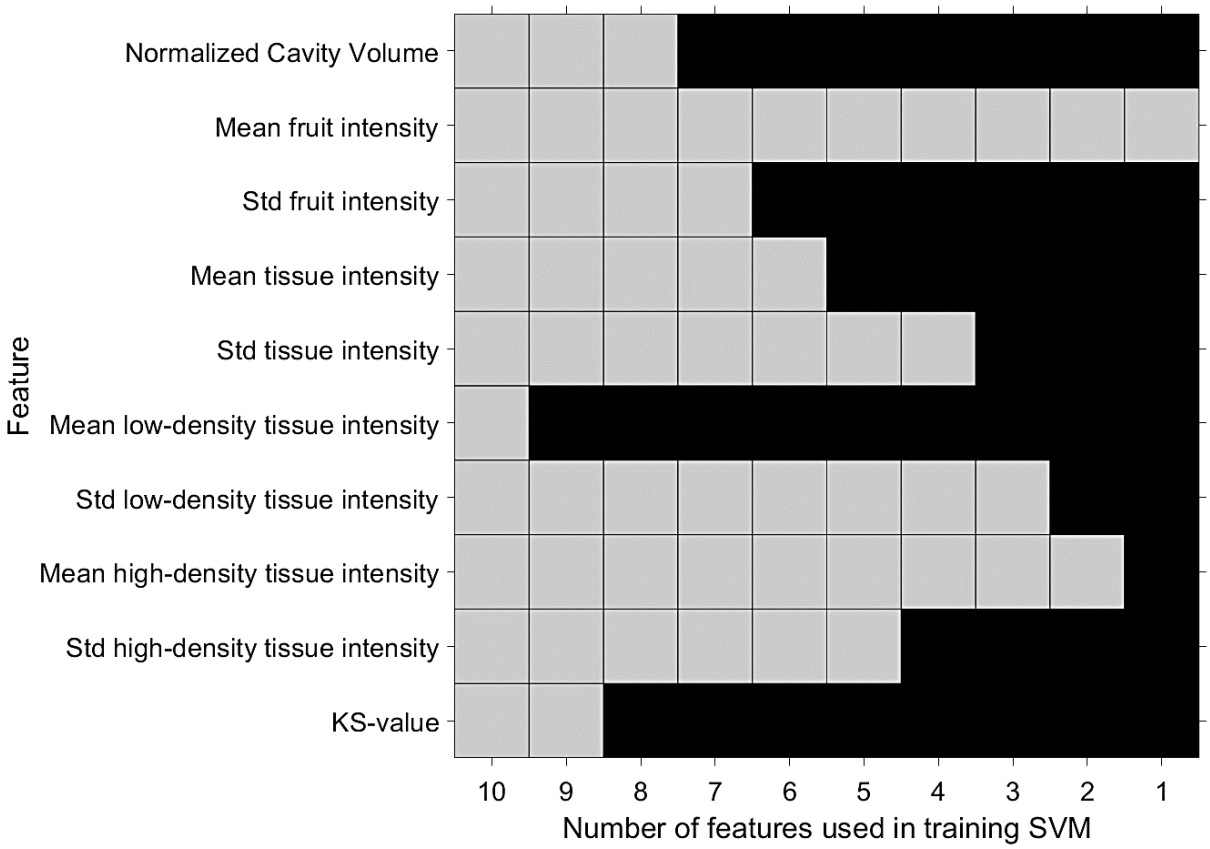

Figure 9: Plot of features used for retraining each SVM on the combined dataset. Each column represents a classifier in a series of classifiers, in which the number of features allowed to be used decreased from the left (10 features) to the right (1 feature). Every column thus shows the features used by a classifier in the series, while each row shows in which classifiers a certain feature was used. A grey tile indicates that the feature was used, while a black tile indicates that a feature was eliminated for the specific SVM in the feature elimination procedure.

4 DISCUSSION

4.1 Internal variability must be measured to separate *'defective'* from *'healthy'* pear fruit

Internal browning and cavity formation were introduced in 'Conference' and 'Cepuna' pears by exposing them to suboptimal storage treatments for six months. The internal disorders differed in severity, location and appearance. Internal browning was characterized by reduced voxel intensity in the CT reconstructions of the fruit due to reduction in tissue density associated with water loss in the affected regions. In regions with severe internal disorder development, cells broke down completely and cavities were observed. This is in line with observations for pear made by (Lammertyn et al., 2003a; Muziri et al., 2016; van Dael et al., 2017).

To further characterize *'defective'* and *'healthy'* pear fruit, features were extracted from the CT volumes and compared between 'Conference' and 'Cepuna' *'defective'* and *'healthy'* fruit (see Table 1). First, the features *'Std fruit intensity'*, *'Std tissue intensity'*, *'Std low-density tissue intensity*' and *'Std high-density tissue intensity*' seemed to be the most relevant ones for separating the classes *'defective'* and *'healthy'*, regardless of the fruit cultivar. The *'defective'* class had significantly higher values for

these four features due to a wider range in voxel intensity and thus in tissue density. This suggests that when looking for features that separate the *'defective'* from *'healthy'* pear fruit, features that measure variability rather than absolute values will be more performant for classifying pear fruit regardless of their cultivar.

Second, for the 'Cepuna' cultivar a significantly lower *'Mean tissue intensity*' for both *'healthy'* and *'defective'* fruit was observed compared to 'Conference' fruit. This might indicate that on average, the 'Cepuna' fruit have a lower density, i.e. higher porosity, than 'Conference' fruit (Nugraha et al., 2019).

Third, other features like *'Mean fruit intensity*' and *'Mean low-density tissue intensity'* were significantly different between the classes for each cultivar, but no clear threshold can be indicated that works for both cultivars.

Fourth, the KS-value showed to be a good feature to separate *'healthy'* and *'defective'* 'Conference' fruit with higher KS-values for *'defective'* fruit. However, the KS-value was not significantly different for both classes of 'Cepuna' fruit. Even an opposite, although not significant, trend was observed with slightly higher values for *'healthy'* fruit.

Finally, compared to the observed *'healthy'* 'Cepuna' fruit, the *'healthy'* 'Conference' fruit had significantly higher normalized cavity volumes. As such, relative to the total fruit volume, 'Conference' fruit might have larger cores than 'Cepuna' fruit.

One must be careful to generalize these results because environmental factors potentially influencing the fruit characteristics were not investigated. Fruit characteristics and susceptibility for internal disorders can be seasonal and location specific. Moreover, only fifteen *'healthy'* 'Cepuna' fruit occurred in the dataset. Unfortunately, a large part of the 'Cepuna' fruit following the control treatment also developed internal disorders. Due to the small sample size of *'healthy'* 'Cepuna' fruit, the observed differences between classes and cultivars must thus be interpreted with caution.

4.2 X-ray CT and machine learning can be implemented inline to classify fruit reliably

The large variability in severity, location and appearance of the internal disorders makes it challenging to develop algorithms that detect *'defective'* fruit reliably. However, for the internal disorder detection in 'Conference' pears, a SVM achieved a classification accuracy of 92.2 % with false positive and false negative rates of respectively 2.4 and 11.7 % (see Table 2). Moreover, the number of features was reduced from ten to two while keeping classification performance high by using the SVM RFE method (see Figure 7 and Figure 8). The classifier trained with the features *'Std tissue intensity'* and *'Std low-density tissue intensity'* still achieved an accuracy, false positive rate and false negative rate of respectively 91.2, 4.8 and 13.3 %.

Furthermore, without retraining or other adaptions to the method the *'Conference' based SVMs'* performed excellent on the 'Cepuna' cultivar as well. An overall classification accuracy of 95.1 % with a false positive and a false negative rate of respectively 0.0 and 5.7 %, was achieved by the ten-feature *'Conference' based SVM'* (see Table 3). Compared to the ten-feature classifier, the two-feature classifier scored the same accuracy with a better false negative rate of 4.6 %, but worse false positive rate of 6.7 %. This shows that the classifiers trained on the 'Conference' cultivar generalize well to the 'Cepuna' cultivar and suggests that the method can be used for other pear cultivars too without much effort. However, an increase in generalizability by reducing the number of features used by the classifiers was not observed, as the performance of all classifiers using between ten and two features was very similar for both cultivars.

The *'Conference' based SVMs'* were compared to two series retrained on the combined dataset. The first series was forced to use the same features as the *'Conference' based SVMs'*, but was allowed to adapt the weights. In the second series, also the selected features were allowed to be altered by re-implementing the SVM RFE method. In both cases, the *'Conference' based SVMs'* scored better, even though no 'Cepuna' data was included in the training process. The two-feature SVM of the second series trained on the combined dataset used the features *'Mean fruit intensity'* and *'Mean high-density tissue intensity'* in contrast to the two-feature *'Conference' based SVM'* that used the features *'Std tissue intensity'* and *'Std low-density tissue intensity'*. Differences in performance and selected features are probably caused by the imbalance in the combined dataset. Only 28 % of the combined dataset was *'healthy'* as just fifteen out of the 102 'Cepuna' samples were *'healthy'*, compared to 42 out of the 102 'Conference' samples. The classification metrics for both two-features classifiers are shown on the combined dataset, the 'Conference' subset and the 'Cepuna' subset in Table 5. The classifier retrained on the combined dataset scored very poorly on *'healthy'* 'Cepuna' fruit, with a 40.0 % false positive rate. Due to the small number of *'healthy'* 'Cepuna' samples, however, this only had a limited effect on the overall classification accuracy over the 'Cepuna' and combined datasets.

Table 5: Confusion matrix with average classification results on the combined dataset, 'Conference' dataset and 'Cepuna' dataset of the two-features 'Conference' based SVM' vs the two-feature classifier that was retrained on the combined dataset with the features 'Mean fruit intensity' and 'Mean high-density tissue intensity'.

Combined dataset		Predicted			
		'Conference' based SVM		Retrained SVM	
		Defective	Healthy	Defective	Healthy
Ground truth	Defective	91.8 %	8.2 %	92.5 %	7.5 %
	Healthy	3.5 %	96.5 %	12.3 %	87.7 %
	Overall classification accuracy:	93.1 %		91.2 %	

'Conference' subset		Predicted			
		'Conference' based SVM		Retrained SVM	
		Defective	Healthy	Defective	Healthy
Ground truth	Defective	86.7 %	13.3 %	85.0 %	15.0 %
	Healthy	4.8 %	95.2 %	4.8 %	95.2 %
	Overall classification accuracy:	91.2 %		89.2 %	

'Cepuna' subset		Predicted			
		'Conference' based SVM		Retrained SVM	
		Defective	Healthy	Defective	Healthy
Ground truth	Defective	95.3 %	4.7 %	97.7 %	2.3 %
	Healthy	6.7 %	93.7 %	40.0 %	60.0 %
	Overall classification accuracy:	95.1 %		92.2 %	

Overall low false positive rates by the *'Conference' based SVMs'* were achieved, ranging between 0.0 and 6.7 % (see Table 2, Table 3, Table 4 and Table 5). Low false positive rates ensure that the number of *'healthy'* fruit that are falsely rejected, are minimized. Furthermore, there were false positives that did not have a pronounced internal disorder but showed small deviating characteristics. For instance, one rejected 'Conference' fruit had a relatively big open core which might be indeed disliked by some consumers (see Figure 10). Economically, it makes sense to minimize the false positive rate for this application, since during inspection the occurrence of internal disorders in a certain batch might be relatively low and the false negatives might not have severe defects. In contrast, with a high false positive rate to ensure a low false negative rate, the added benefit of detecting *'defective'* fruit might be offset by the falsely rejected *'healthy'* fruit. Of course, this depends on the severity of the internal disorders. To balance the compromise between the false positive and false negative rate in a desired way, one could set a different threshold for the decision boundary of the classifier instead of placing it at f(x) = 0. For 'Conference', the classifier had a true positive rate of 86.7 % with a false positive rate of

4.8 %. To have 0 % false positives, the true positive rate had to drop to around 82 %. To achieve a true positive rate of 100 %, the false positive rate increased to 60 %.

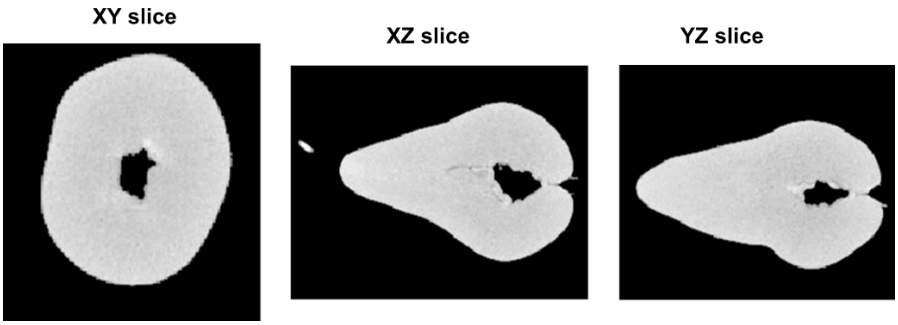

Figure 10: Orthogonal slices through the CT reconstructions of a false positive 'Conference' pear example with relatively big open core.

The *'Conference' based SVMs'* had false negative rates ranging between 11.7 and 13.3 % for 'Conference' and between 4.7 and 5.7 % for 'Cepuna'. Compared to the false positive rates, the false negative rates are thus higher. As explained above, however, in this application it might be more important to reduce the false positive rate. Moreover, most of the false negative samples had only a very small internal defects that may not even be noted by the end consumer. As an example, Figure 11 shows orthogonal slices through the CT reconstructions of a 'Conference' and 'Cepuna' pear. These examples make clear that for interpreting classification results, it is important to investigate how the data was labeled. To consider the consumer acceptance and preferences in future research, it is suggested to perform an expert panel survey for labeling the fruit based on images of cut fruit in addition to a visual inspection of the CT data.

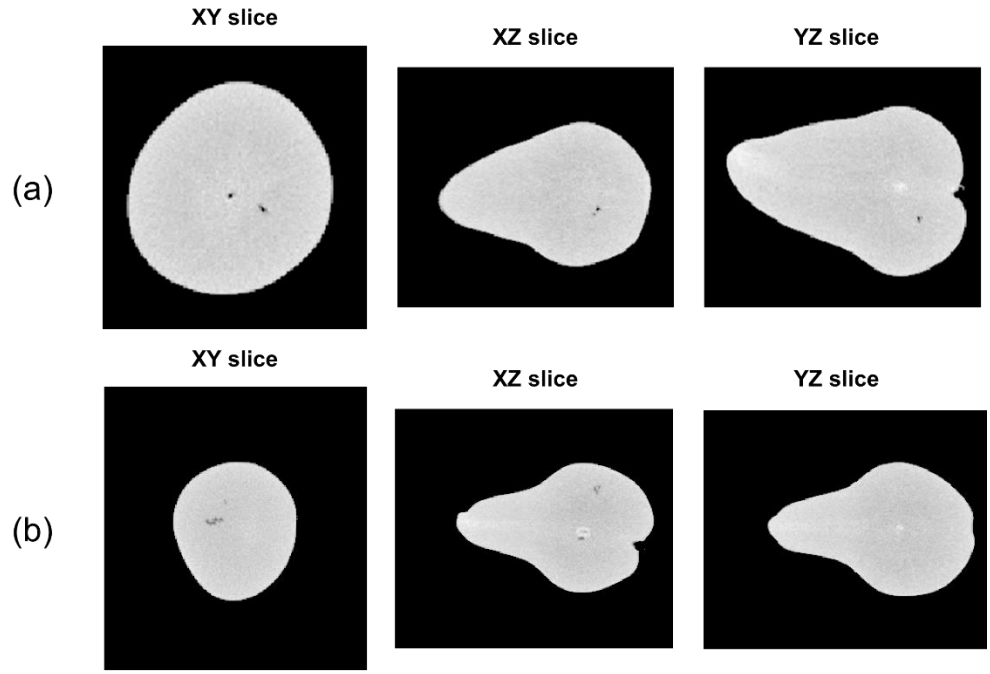

Figure 11: Orthogonal slices through the CT reconstructions of a false negative 'Conference' (a) and 'Cepuna' (b) pear examples.

This analysis showed that the inspection of the internal quality of pear fruit can be done in 3D using X-ray CT at high classification accuracies and low false positive rates. In the current experiment, the achieved inspection speed (1.5 minute for scanning, stacking and reconstruction, plus 2.3 s for feature extraction and classification) was not yet compatible with the commercial speed of existing sorting lines, requiring at least 10 fruit per second per lane. Nonetheless, it should be noted that both the experimental CT setup and the feature extraction algorithm were not optimized for reducing the runtime. With advanced reconstruction algorithms available for translational X-ray CT (De Schryver et al., 2016; Janssens et al., 2018), prototypes of inline CT systems can be developed. Speed improvements can be made using a dedicated system combined with a trained reconstruction algorithm that needs far less projections (Beister et al., 2012; De Schryver et al., 2016; Janssens et al., 2018; Willemink et al., 2013). Moreover, a stacked scan was needed due to the relatively small detector size, requiring two full 360° sample rotations, intermediate height adjustments and a stacking procedure. With a detector of appropriate size, the scanning time can thus already be reduced by 50 %. The method could also be tested with other scanning settings, like exposure time and number of projections to reduce the scanning time, or pixel and voxel sizes to reduce the computational cost during reconstruction and analysis. Additionally, the feature extraction algorithm could be optimized and it was shown that the number of features can be reduced to further reduce the processing time. In terms of hardware, technical challenges must be overcome to transition from prototype systems to fast inline CT systems, *e.g.,* the development sample holders that stabilize and rotate the sample while translating at adequate speed. For industrial application continuing developments in both hardware

and software are thus needed to increase the inspection speed and reduce the equipment costs. In addition, the internal quality inspection system can be used on a limited part of the supply in which a high occurrence of internal disorders is expected. In this case, lower inspection speeds might be adequate.

The proposed method uses machine learning in which a classifier is trained on a training set in feature space. However, the features were still hand-crafted and thus possibly application specific. Application on other fruit products, like apples, might require other features to be used making feature determination more difficult. In the feature extraction algorithm, the KS-value was calculated as a measure of similarity between the low intensity (low density) and high intensity (high density) tissue regions. It was expected to be an important feature for separating the classes, however only low corresponding weights were given to this feature by the classifiers. This illustrates that intuitive hand-crafted features that seem smartly designed, may not always be the best choice, and presents a limitation of machine learning. In deep learning, valuable representations of the data are learned and extracted by the model itself. Hand-crafted features must thus no longer be engineered (Goodfellow et al., 2016). Deep learning on X-ray imaging is increasingly adopted, e.g., in medical applications (Lee et al., 2017; Shen et al., 2017), and might thus be considered as an alternative method in future research.

5 CONCLUSION

A combination of machine learning and X-ray computed tomography was proposed to successfully detect 'Conference' and 'Cepuna' pear fruit with a wide range in internal disorder severity automatically. Trained SVMs achieved good classification accuracies ranging between 90.2 and 95.1 % depending on the cultivar and number of features that were used. Moreover, low false positive and negative rates were obtained, respectively ranging between 0.0 and 6.7 %, and 5.7 and 13.3 %. Classifiers trained on 'Conference' data achieved high validation scores on the 'Cepuna' cultivar suggesting generalizability to other cultivars as well.

With continuing developments in both hardware and software to increase inspection speed and reduce the equipment costs, the method can be implemented in *e.g.,* inline translational X-ray CT for industrial application. Further research should focus on testing the method on an inline translational CT system with various parameter settings for image acquisition, validating the method for other cultivars or fruit and investigate other methods, *e.g.,* deep learning, for quality inspection of food in 3D X-ray CT imaging.

6 DECLARATION OF COMPETING INTEREST

None.

7 ACKNOWLEDGEMENTS

This research was funded by the imec ICON project iXCon (Agentschap Innoveren & Ondernemen project nr. HBC.2016.0164), the Flanders' Food i-FAST project (Agentschap Innoveren & Ondernemen project nr. IWT 140992) and the iProcess project (Research Council of Norway project nr. 255596/E50).

8 REFERENCES

Alves Pereira, L.F., Janssens, E., Cavalcanti, G.D.C., Tsang, I.R., Van Dael, M., Verboven, P., Nicolai, B., Sijbers, J., 2017. Inline discrete tomography system: Application to agricultural product inspection. Comput. Electron. Agric. 138, 117–126. https://doi.org/10.1016/J.COMPAG.2017.04.010

Arendse, E., Fawole, O.A., Magwaza, L.S., Opara, U.L., 2018. Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. J. Food Eng. https://doi.org/10.1016/j.jfoodeng.2017.08.009

Beister, M., Kolditz, D., Kalender, W.A., 2012. Iterative reconstruction methods in X-ray CT. Phys. Medica 28, 94–108. https://doi.org/10.1016/j.ejmp.2012.01.003

Bobelyn, E., Serban, A.-S., Nicu, M., Lammertyn, J., Nicolai, B.M., Saeys, W., 2010. Postharvest quality of apple predicted by NIR-spectroscopy: Study of the effect of biological variability on spectra and model performance. Postharvest Biol. Technol. 55, 133–143. https://doi.org/10.1016/J.POSTHARVBIO.2009.09.006

Buratti, A., Bredemann, J., Pavan, M., Schmitt, R., Carmignato, S., 2018. Applications of CT for Dimensional Metrology, in: Carmignato, S., Dewulf, W., Leach, R. (Eds.), Industrial X-Ray Computed Tomography. Springer International Publishing, Cham, pp. 333–369. https://doi.org/10.1007/978-3-319-59573-3_9

Casasent, D.A., Sipe, M.A., Schatzki, T.F., Keagy, P.M., Lee, L.C., 1998. Neural net classification of X-ray pistachio nut data. LWT - Food Sci. Technol. 31, 122–128. https://doi.org/10.1006/fstl.1997.0320

Colnago, L.A., Andrade, F.D., Souza, A.A., Azeredo, R.B.V., Lima, A.A., Cerioni, L.M., Osán, T.M., Pusiol, D.J., 2014. Why is Inline NMR Rarely Used as Industrial Sensor? Challenges and Opportunities. Chem. Eng. Technol. 37, 191–203. https://doi.org/10.1002/ceat.201300380

De Schryver, T., Dhaene, J., Dierick, M., Boone, M.N., Janssens, E., Sijbers, J., Dael, M. Van,

569 Verboven, P., Nicolai, B., 2016. In-line NDT with X-Ray CT combining sample rotation and
570 translation. NDT E Int. 84, 89–98. https://doi.org/10.1016/j.ndteint.2016.09.001

571 Donis-González, I.R., Guyer, D.E., Pease, A., 2016a. Postharvest noninvasive classification of tough-
572 fibrous asparagus using computed tomography images. Postharvest Biol. Technol. 121, 27–35.
573 https://doi.org/10.1016/J.POSTHARVBIO.2016.07.012

574 Donis-González, I.R., Guyer, D.E., Pease, A., 2016b. Postharvest noninvasive assessment of undesirable
575 fibrous tissue in fresh processing carrots using computer tomography images. J. Food Eng. 190,
576 154–166. https://doi.org/10.1016/J.JFOODENG.2016.06.024

577 Donis-González, I.R., Guyer, D.E., Pease, A., Barthel, F., 2014. Internal characterisation of fresh
578 agricultural products using traditional and ultrafast electron beam X-ray computed tomography
579 imaging. Biosyst. Eng. 117, 104–113. https://doi.org/10.1016/J.BIOSYSTEMSENG.2013.07.002

580 Franck, C., Lammertyn, J., Ho, Q.T., Verboven, P., Verlinden, B., Nicolaï, B.M., 2007. Browning
581 disorders in pear fruit. Postharvest Biol. Technol. 43, 1–13.
582 https://doi.org/10.1016/j.postharvbio.2006.08.008

583 Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep learning, Adaptive computation and machine
584 learning. The MIT Press, Cambridge, Massachusetts.

585 Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using
586 support vector machines. Mach. Learn. 46, 389–422. https://doi.org/10.1023/A:1012487302797

587 Han, D., Tu, R., Lu, C., Liu, X., Wen, Z., 2006. Nondestructive detection of brown core in the Chinese
588 pear 'Yali' by transmission visible–NIR spectroscopy. Food Control 17, 604–608.
589 https://doi.org/10.1016/J.FOODCONT.2005.03.006

590 Herremans, E., Melado-Herreros, A., Defraeye, T., Verlinden, B., Hertog, M., Verboven, P., Val, J.,
591 Fernández-Valle, M.E., Bongaers, E., Estrade, P., Wevers, M., Barreiro, P., Nicolaï, B.M., 2014.
592 Comparison of X-ray CT and MRI of watercore disorder of different apple cultivars. Postharvest
593 Biol. Technol. https://doi.org/10.1016/j.postharvbio.2013.08.008

594 Herremans, E., Verboven, P., Bongaers, E., Estrade, P., Verlinden, B.E., Wevers, M., Hertog,
595 M.L.A.T.M., Nicolai, B.M., 2013. Characterisation of "Braeburn" browning disorder by means of
596 X-ray micro-CT. Postharvest Biol. Technol. https://doi.org/10.1016/j.postharvbio.2012.08.008

597 Huang, Y., Lu, R., Chen, K., 2020. Detection of internal defect of apples by a multichannel Vis/NIR
598 spectroscopic system. Postharvest Biol. Technol. 161, 111065.
599 https://doi.org/10.1016/J.POSTHARVBIO.2019.111065

600 Janssens, E., Alves Pereira, L.F., De Beenhouwer, J., Tsang, I.R., Van Dael, M., Verboven, P., Nicolaï,
601 B., Sijbers, J., 2016. Fast inline inspection by Neural Network Based Filtered Backprojection:

Application to apple inspection. Case Stud. Nondestruct. Test. Eval. 6, 14–20. https://doi.org/10.1016/j.csndt.2016.03.003

Janssens, E., De Beenhouwer, J., Van Dael, M., De Schryver, T., Van Hoorebeke, L., Verboven, P., Nicolai, B., Sijbers, J., 2018. Neural network Hilbert transform based filtered backprojection for fast inline x-ray inspection. Meas. Sci. Technol. 29. https://doi.org/10.1088/1361-6501/aa9de3

Jarolmasjed, S., Espinoza, C.Z., Sankaran, S., Khot, L.R., 2016. Postharvest bitter pit detection and progression evaluation in 'Honeycrisp' apples using computed tomography images. Postharvest Biol. Technol. 118, 35–42. https://doi.org/10.1016/J.POSTHARVBIO.2016.03.014

Jiang, J.A., Chang, H.Y., Wu, K.H., Ouyang, C.S., Yang, M.M., Yang, E.C., Chen, T.W., Lin, T. Te, 2008. An adaptive image segmentation algorithm for X-ray quarantine inspection of selected fruits. Comput. Electron. Agric. 60, 190–200. https://doi.org/10.1016/j.compag.2007.08.006

Karunakaran, C., Jayas, D.S., White, N.D.G., 2004. Identification of wheat kernels damaged by the red flour beetle using X-ray images. Biosyst. Eng. 87, 267–274. https://doi.org/10.1016/j.biosystemseng.2003.12.002

Khatiwada, B.P., Subedi, P.P., Hayes, C., Carlos, L.C.C., Walsh, K.B., 2016. Assessment of internal flesh browning in intact apple using visible-short wave near infrared spectroscopy. Postharvest Biol. Technol. 120, 103–111. https://doi.org/10.1016/J.POSTHARVBIO.2016.06.001

Kim, S., Schatzki, T., 2001. Detection of Pinholes in Almonds through X-ray. Trans. ASAE 44, 997–1003.

Kim, S., Schatzki, T., 2000. Apple watercore sorting system using X-ray imagery: I. Algorithm development. Trans. Am. Soc. Agric. Biol. Eng. 43, 1695–1702.

Kotwaliwale, N., Singh, K., Kalne, A., Jha, S.N., Seth, N., Kar, A., 2014. X-ray imaging methods for internal quality evaluation of agricultural produce. J. Food Sci. Technol. https://doi.org/10.1007/s13197-011-0485-y

Kotwaliwale, N., Weckler, P.R., Brusewitz, G.H., Kranzler, G.A., Maness, N.O., 2007. Non-destructive quality determination of pecans using soft X-rays. Postharvest Biol. Technol. 45, 372–380. https://doi.org/10.1016/J.POSTHARVBIO.2007.03.008

Lammertyn, J., Aerts, M., Verlinden, B.E., Schotsmans, W., Nicolaï, B.M., 2000. Logistic regression analysis of factors influencing core breakdown in "Conference" pears. Postharvest Biol. Technol. 20, 25–37. https://doi.org/10.1016/S0925-5214(00)00114-9

Lammertyn, J., Dresselaers, T., Van Hecke, P., Jancsók, P., Wevers, M., Nicolaï, B.M., 2003a. MRI and X-ray CT study of spatial distribution of core breakdown in "Conference" pears. Magn. Reson. Imaging 21, 805–815. https://doi.org/10.1016/S0730-725X(03)00105-X

Lammertyn, J., Dresselaers, T., Van Hecke, P., Jancsók, P., Wevers, M., Nicolaï, B.M., 2003b. Analysis of the time course of core breakdown in “Conference” pears by means of MRI and X-ray CT. Postharvest Biol. Technol. 29, 19–28. https://doi.org/10.1016/S0925-5214(02)00212-0

Lee, J.G., Jun, S., Cho, Y.W., Lee, H., Kim, G.B., Seo, J.B., Kim, N., 2017. Deep learning in medical imaging: General overview. Korean J. Radiol. 18, 570–584. https://doi.org/10.3348/kjr.2017.18.4.570

Lu, Y., Lu, R., 2017. Non-Destructive Defect Detection of Apples by Spectroscopic and Imaging Technologies: A Review. Trans. ASABE 60, 1765. https://doi.org/https://doi.org/10.13031/trans.12431

Massey, F.J.J., 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. J. Am. Stat. Assoc. 46, 68–78.

MATLAB, 2019a. MATLAB.

MATLAB, 2019b. Two-sample Kolmogorov-Smirnov test [WWW Document]. URL https://nl.mathworks.com/help/stats/kstest2.html#btn37ur (accessed 7.1.19).

Mercier, S., Villeneuve, S., Mondor, M., Uysal, I., 2017. Time–Temperature Management Along the Food Cold Chain: A Review of Recent Developments. Compr. Rev. Food Sci. Food Saf. 16, 647–667. https://doi.org/10.1111/1541-4337.12269

Muziri, T., Theron, K.I., Cantre, D., Wang, Z., Verboven, P., Nicolai, B.M., Crouch, E.M., 2016. Microstructure analysis and detection of mealiness in ‘Forelle’ pear (Pyrus communis L.) by means of X-ray computed tomography. Postharvest Biol. Technol. 120, 145–156. https://doi.org/10.1016/j.postharvbio.2016.06.006

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biol. Technol. 46, 99–118. https://doi.org/10.1016/J.POSTHARVBIO.2007.06.024

Nicolaï, B.M., Defraeye, T., De Ketelaere, B., Herremans, E., Hertog, M.L.A.T.M., Saeys, W., Torricelli, A., Vandendriessche, T., Verboven, P., 2014. Nondestructive Measurement of Fruit and Vegetable Quality. Annu. Rev. Food Sci. Technol. https://doi.org/10.1146/annurev-food-030713-092410

Nugraha, B., Verboven, P., Janssen, S., Wang, Z., Nicolaï, B.M., 2019. Non-destructive porosity mapping of fruit and vegetables using X-ray CT. Postharvest Biol. Technol. 150, 80–88. https://doi.org/10.1016/J.POSTHARVBIO.2018.12.016

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man.

668 Cybern. 9, 62–66. https://doi.org/10.1109/TSMC.1979.4310076

669 Pedreschi, R., Franck, C., Lammertyn, J., Erban, A., Kopka, J., Hertog, M., Verlinden, B., Nicolaï, B.,
670 2009. Metabolic profiling of “Conference” pears under low oxygen stress. Postharvest Biol.
671 Technol. 51, 123–130. https://doi.org/10.1016/j.postharvbio.2008.05.019

672 Shahin, M.A., Tollner, E.W., McClendon, R.W., 2001. Artificial intelligence classifiers for sorting
673 apples based on watercore. J. Agric. Eng. Res. 79, 265–274.
674 https://doi.org/10.1006/jaer.2001.0705

675 Shen, D., Wu, G., Suk, H.-I., 2017. Deep Learning in Medical Image Analysis. Annu. Rev. Biomed.
676 Eng. 19, 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

677 Si, Y., Sankaran, S., 2016. Computed tomography imaging-based bitter pit evaluation in apples. Biosyst.
678 Eng. 151, 9–16. https://doi.org/10.1016/J.BIOSYSTEMSENG.2016.08.008

679 Srivastava, R.K., Talluri, S., Khasim Beebi, S., Kumar, R., 2018. Magnetic Resonance Imaging for
680 Quality Evaluation of Fruits: a Review. Food Anal. Methods 11, 2943–2960.
681 https://doi.org/10.1007/s12161-018-1262-6

682 Statbel, 2018. Tab A landbouwcijfers 2018 [WWW Document]. Landbouwgegevens van 2018. URL
683 https://statbel.fgov.be/sites/default/files/files/documents/landbouw/8.1 Land- en
684 tuinbouwbedrijven/DBREF-L05-2018-TAB-A-NL.xlsx

685 van Dael, M., Lebotsa, S., Herremans, E., Verboven, P., Sijbers, J., Opara, U.L., Cronje, P.J., Nicolaï,
686 B.M., 2016. A segmentation and classification algorithm for online detection of internal disorders
687 in citrus using X-ray radiographs. Postharvest Biol. Technol.
688 https://doi.org/10.1016/j.postharvbio.2015.09.020

689 van Dael, M., Rogge, S., Verboven, P., Saeys, W., Sijbers, J., Nicolai, B., 2015. Online Tomato
690 Inspection Using X-Ray Radiographies and 3- Dimensional Shape Models. Chem. Eng. Trans. 44.

691 van Dael, M., Verboven, P., Dhaene, J., Van Hoorebeke, L., Sijbers, J., Nicolai, B., 2017. Multisensor
692 X-ray inspection of internal defects in horticultural products. Postharvest Biol. Technol. 128, 33–
693 43. https://doi.org/10.1016/j.postharvbio.2017.02.002

694 van Dael, M., Verboven, P., Zanella, A., Sijbers, J., Nicolai, B., 2018. Combination of shape and X-ray
695 inspection for apple internal quality control: in silico analysis of the methodology based on X-ray
696 computed tomography. Postharvest Biol. Technol. 0–1.
697 https://doi.org/10.1016/j.postharvbio.2018.05.020

698 VCBT, 2017. Bewaarcondities Appel en Peer [WWW Document]. URL
699 http://vcbt.be/bewaarcondities_appel_en_peer/ (accessed 9.1.17).

700 Veltman, R.H., Lenthéric, I., Van Der Plas, L.H.W., Peppelenbos, H.W., 2003. Internal browning in

pear fruit (Pyrus communis L. cv Conference) may be a result of a limited availability of energy and antioxidants. Postharvest Biol. Technol. 28, 295–302. https://doi.org/10.1016/S0925-5214(02)00198-9

Wang, Z., Herremans, E., Janssen, S., Cantre, D., Verboven, P., Nicolaï, B., 2018. Visualizing 3D Food Microstructure Using Tomographic Methods: Advantages and Disadvantages. Annu. Rev. Food Sci. Technol. 9, 323–343. https://doi.org/10.1146/annurev-food-030117-012639

Wevers, M., Nicolaï, B., Verboven, P., Swennen, R., Roels, S., Verstrynge, E., Lomov, S., Kerckhofs, G., Van Meerbeek, B., Mavridou, A.M., Bergmans, L., Lambrechts, P., Soete, J., Claes, S., Claes, H., 2018. Applications of CT for Non-destructive Testing and Materials Characterization, in: Carmignato, S., Dewulf, W., Leach, R. (Eds.), Industrial X-Ray Computed Tomography. Springer International Publishing, Cham, pp. 267–331. https://doi.org/10.1007/978-3-319-59573-3_8

Willemink, M.J., De Jong, P.A., Leiner, T., De Heer, L.M., Nievelstein, R.A.J., Budde, R.P.J., Schilham, A.M.R., 2013. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur. Radiol. 23, 1623–1631. https://doi.org/10.1007/s00330-012-2765-y