Non-destructive internal disorder segmentation in pear fruit by X-ray radiography and AI

Astrid Tempelaere[1,a], Hoang Minh Phan[1,b], Tim Van De Looverbosch[2,c], Pieter Verboven[1,d,*], Bart Nicolai[1,3,e]

1 Mechatronics, Biostatistics and Sensors (MeBioS), Biosystems Department, KU Leuven, Belgium

2 Laboratory of Cell Biology and Histology, UAntwerpen, Belgium

3 Flanders Centre of Postharvest Technology, Belgium

[a]astrid.tempelaere@kuleuven.be; [b]hoangminh.phan@kuleuven.be;

[c]tim.vandelooverbosch@uantwerpen.be; [d]pieter.verboven@kuleuven.be;

[e]bart.nicolai@kuleuven.be

*Corresponding author

Abstract

A particular challenge for food quality inspection remains the quantification of internal defects. This work addresses the challenge of internal defects segmentation in pear fruit (cv. Conference) based on high-throughput inline X-ray radiography images in combination with deep learning. Unlike previous approaches, which focused on healthy vs defect classification, the current segmentation approach is able to determine the type of defect, its dimensions, and location. To this end, a novel simulation method was designed to obtain input-target image pairs of radiography data, and more diverse defect pears were generated using a conditioned generator model. The obtained data contributed to the design of a segmentation model that labeled every pixel in the X-ray radiographs of pear fruit as 'external air', 'healthy tissue', 'core', 'browning', or 'cavity'. We demonstrated that additional synthetic data in the learning process drastically improved the model performance. For instance, the mean IoU increased from 0.781 ± 0.112 to 0.883 ± 0.088 for consumable pears with minor cavities and browning. In terms of utility, our segmentation maps provide more detailed information about the

type, size, and location of the disorders compared to the heatmaps produced by existing benchmark classifiers.

1. Introduction

To maintain fruit quality after harvest and meet the year-round demand, pome fruit is commonly stored under controlled atmosphere conditions (Watkins et al., 2020). However, internal disorders may develop due to various preharvest factors and storage conditions, leading to significant losses. These disorders are externally invisible and typically involve internal tissue browning and cavity formation (Franck et al., 2007). X-ray imaging is a promising technology for non-destructive internal quality inspection of fruit as it provides information on differences in tissue density (Nicolaï et al., 2014). X-ray imaging research for internal quality inspection of fruit has mainly focused on X-ray computed tomography (CT) (Barcelon et al., 1999; Cantre et al., 2014; Herremans et al., 2014; Janssens et al., 2016; Ogawa et al., 2003; Poly et al., 2018; Van De Looverbosch et al., 2021, 2020). This technique provides three-dimensional (3D) image data on the internal structure of the inspected object. High classification accuracies on X-ray CT images of pear fruit have been demonstrated using artificial intelligence (AI) (Van De Looverbosch et al., 2021, 2020). However, several issues hamper its industrial implementation. Although it has been shown in laboratory conditions that in-line X-ray CT imaging is feasible by combining sample rotation and translation, it is difficult to achieve vibrationless transport of the fruit in industrial sorting lines (De Schryver et al., 2016). Such unplanned movements may cause motion artefacts in the reconstructed images or even impede image reconstruction. In addition, commercial sorting lines typically operate at ten pieces of fruit per second, leaving less than 100 ms for image construction. Although fast alternatives to the common filtered back projection algorithms that also can deal with a smaller number of radiographies have been developed, achieving fast image reconstruction and further processing them is still challenging (Alves Pereira et al., 2017; Janssens et al., 2016). Finally, X-ray CT is expensive.

To avoid the practical shortcomings of X-ray CT, X-ray radiography has been explored as an alternative technology in which the 3D object is mapped to a two-dimensional (2D) projection, i.e., a radiograph, which is faster to acquire with more simple equipment compatible with sorting lines (Nicolaï et al., 2014). However, this comes at the expense of losing some spatial information and contrast in the image, rendering image analysis more difficult (Belin et al., 2011; Finney and Norris, 1978; Gadgile et al., 2017; Gadgile and Chavan, 2017; Medeiros et al., 2020; Rady et al., 2020; Schatzki et al., 1996; Shahin et al., 2002, 2001; Sherif et al., 2023; van Dael et al., 2016; S. Yu et al., 2022).

The possibility of using X-ray radiography for internal quality inspection of horticultural products was described for the first time in a study about hollow heart in potatoes (Finney and Norris, 1978). Also, Schatzki et al. (1996) explored this technology for the detection of internal disorders in apple fruit. More recently, other researchers reported that fungal diseases in mango and citrus fruit can be detected by this technology (Gadgile et al., 2017; Gadgile and Chavan, 2017). Another study also demonstrated that freezing injury in pear fruit can be observed in the radiographs (S. Yu et al., 2022). Initially, the X-ray data were evaluated by visual observation (Gadgile et al., 2017; Gadgile and Chavan, 2017; Schatzki et al., 1996). To automate this process, Finney & Norris (1978) developed a classification algorithm for X-ray images on hollow heart in potatoes by observing the X-ray density profile along the longitudinal axis. Machine learning (ML) classifiers were also defined to automatically detect internal quality using preselected image features. In this context, prior works were found on bruised or watercore apples (Shahin et al., 2002, 2001), citrus fruit with granulation or endoxerosis (van Dael et al., 2016), fasciation in lettuce (Sherif et al., 2023), and seed quality (Ahmed et al., 2020; Medeiros et al., 2020). The features used in ML are typically determined by the experts themselves, which requires careful analysis and selection to render robust and reliable classifiers.

In the more recent field of deep learning (DL), the model itself decides the features that explain the data the best (Goodfellow et al., 2016). Earlier studies developed DL models for internal quality detection of seeds and cereals based on radiographs (Ahmed et al., 2020; da Silva et al., 2021; Medeiros et al., 2021; Xue et al., 2022). However, a large number of ground-truth data is typically required to develop new DL models. In applications of internal disorders of fruit, this is not always possible and typically many more sound fruit are available than defect ones. A recent work approached this task as an anomaly detection problem in which healthy pears were considered as normal and disordered pears as anomalous (Van De Looverbosch et al. 2022). Both unsupervised and supervised approaches were successful, reaching an accuracy of 92% for the former and 96% for the latter method to classify healthy and defect fruit with both cavities and browning. Despite these high numbers, defect pears with only browning were more difficult to distinguish from healthy samples. To gain more insight into the decision-making process of the models, heatmaps were produced that indicate which regions in the radiograph were regarded as anomalous. Comparing the heatmaps with the radiographs demonstrated that the model mostly focused on cavities. In addition to anomaly detection, Tempelaere et al. (2023) addressed the issue of data scarcity by generating synthetic data for healthy and defect pears through a conditional Generative Adversarial Network (cGAN). The synthetic X-ray data from the pears can expand datasets essential for algorithm development in future research.

Notwithstanding high classification successes into healthy and defect fruit, labelling and quantification of internal quality deterioration is particularly interesting to gain better insights into the type, location, and size of disorders in pear by determining the object class on a per-pixel level. In this article, we aim to develop a deep learning segmentation algorithm for improved detection and labeling of the different parts (healthy tissue, brown tissue, cavities and core) of pear fruit based on X-ray radiographs, also called semantic segmentation (Goodfellow et al., 2016). Not only does it provide localization in the image plane, but the semantic information also has a higher utility than the earlier defined heatmaps, as these latter only show the most important regions on which the model bases its decision (Van De Looverbosch et al., 2022). Current state-of-the-art methods for semantic segmentation are dominated by supervised DL, which comes with the requirement of large and detailed annotated datasets. However, manual per-pixel annotations are labor intensive, particularly in domains like agriculture with high environmental and object complexity. In addition, labeling X-ray data of fruit requires expert and domain knowledge. Previous research was able to carefully label 3D X-ray CT data, but argued that manual labeling of 2D radiography data is difficult (Van De Looverbosch et al., 2022). Therefore, this current work introduces a new simulation method to generate input-target image pairs of radiography data based on CT data. As the design of a segmentation model typically requires a large dataset, bootstrapping DL models with synthetic data is a popular approach. The hypothesis is that for effective and successful learning, a small manually annotated empirical dataset combined with related synthetic data would suffice. While our previous work provided initial evidence supporting this hypothesis as a potential avenue for future research, this current study aims to further elaborate on it (Tempelaere et al., 2023). We will build upon our earlier experiments on generating synthetic pear samples and contextualize the results within a broader framework in this paper.

The current study aims to make advances to the application of X-ray imaging for agrofood internal quality inspection by introducing novel image segmentation to identify and localize cavities and browning disorder in pear fruit, while reducing the reliance on labor-intensive manual annotations for training the segmentation algorithms. The main expected contributions of this research will be as follows:

- Synthesis of a wide range of defects in pear fruit for training, using a workflow for generating CT images through a conditional Generative Adversarial Network (cGAN) (Tempelaere et al., 2023).
- Development of a simulation method for input-target image pairs of inline X-ray radiography data, based on the X-ray CT images.

- Design of a deep learning (DL) model for semantic segmentation of healthy and defect pear fruit based on the X-ray radiographs, compared with classification results by state-of-the-art methods (Van De Looverbosch et al. 2022).
- Augmentation of the DL training set by synthetic data, resulting in improved segmentation outcomes.

2. Materials and methods

2.1 X-ray CT data of pear fruit

A CT dataset (8-bit) of 180 pear fruit was acquired by a gantry CT system (SOMATOM Definition, Flash, Siemens, Germany) in the UZ Leuven hospital (Leuven, Belgium) as described previously (Van De Looverbosch et al., 2021). The system operated at 100 kVp with a voxel size of 0.9766 x 0.9766 x 0.300 mm^3. Next, the CT data was processed in MATLAB (2020a, The MathWorks, Natrick, MA, USA). The volumes were resampled to a voxel size of 0.9766 x 0.9766 x 0.9766 mm^3 using cubic interpolation. The background was removed with a global Otsu threshold (Otsu et al., 1979), and the pears were rotated to align their principal axis with the X-axis. Finally, the pear volumes were padded to the same size with background voxels, resulting in a volume of 184 x 179 x 198 voxels for each pear. The voxels in the CT scans from a selection of pears were manually annotated to one of the following five labels: 'external air', 'healthy tissue', 'core', 'cavity', and 'internal browning'. After the CT scans were made, fruit were cut in half and reference images of the fruit flesh were taken. Then, these reference images were shown to participants in a small consumer trial in which the pears were classified as 1) healthy, 2) defect but consumable, or 3) defect and non-consumable. This classification was used to select pears for the training and test sets. Fig. 1 (a) and (b) illustrate a CT slice of a pear which was perceived as 'defect and non-consumable' in grayscale and colorscale, respectively. The corresponding annotated image slice with five labels is shown in Fig. 1C. From the 180 pears, 51 samples (19 healthy, 15 defect and consumable, and 17 defect and non-consumable) were manually annotated and later used as ground-truth data during training and testing (section 2.5). More details on the exact data acquisition, annotation workflow, and consumer test can be found in Van De Looverbosch et al. (2021).

2.2 Simulation method X-ray radiography data

From each pear CT volume, 50 radiographs were simulated by the ASTRA Toolbox (van Aarle et al., 2015) in MATLAB (2020a, The MathWorks, Natick, MA, USA). The radiography simulation required the configuration of the projection geometry. A virtual X-ray line scanner was set with a distance from the X-ray source to the conveyor belt of 0.5 m, and a distance from the conveyor to the detector of 0.05

m to simulate a typical setup that would be used in practice. The speed of the conveyor was set at 0.27 m/s. The detector had a pixel size of 0.5 mm and a width of 300 pixels. The pears were rotated along their principal axis to obtain simulated radiographs from different rotation angles.

Annotated radiographs were simulated by a thresholding method that is illustrated in Fig. 3. The algorithm started from the labeled CT data that were available. For the whole fruit, browning, cavities, and core, binary volumes were extracted. From all these binary volumes, radiographs were simulated and subsequently thresholded by a global threshold of 0, resulting in binary radiographs. Afterwards, these binary images were combined resulting in the labeled radiograph. The core and the cavities were considered to dominate in the radiograph representations compared to browning and healthy tissue. Therefore, an overlay of binary images was performed in a particular order: whole fruit, browning, cavities, and, eventually, core. The labeled radiograph was then used as ground-truth data for the development of our segmentation model (section 2.3).

2.3 Segmentation of pear tissue on X-ray radiography data

2.3.1 U-Net architecture

For automatic segmentation of the radiographs, a U-Net architecture was used, as this model was successful in earlier research on X-ray images (Van De Looverbosch et al., 2021). The original model of Ronneberger et al. (2015) was applied with some small modifications. The input layer consisted of one channel of 300 x 300 pixels. The convolutional layers consisted of 3 x 3 convolutions with batch normalization and a ReLU activation function. The number of downsampling and upsampling blocks were the same as in the original work. However, padding was applied in the convolutional operations to obtain the same dimensions as the data from the corresponding downsampling block. These can then be concatenated in the upsampling path without data loss at the image borders. In addition, bilinear upsampling for fast calculations was used instead of 2 x 2 upconvolutions in the original work. For the output layer, the number of classes was equal to five, corresponding to ‘external air’, ‘healthy tissue’, ‘core’, ‘internal browning’, and ‘cavity’. Finally, the channel with the highest probability at each pixel was obtained in the resulting predicted segmentation. More details on the model architecture are provided in the Supplementary Materials.

2.3.2 Model training & validation

The objective of model training is to lower the loss calculated between the predicted image and the ground-truth. In our work, the Adam optimizer was used to reach this minimum (Kingma and Ba, 2015). More specifically, a cross-entropy loss was applied with class weights to overcome class imbalances (Rahman & Wang, 2016; Zhang & Sabuncu, 2018). The class weights were calculated by the ratio of the pixel number of each class over the number of all the pixels, and these values were subtracted from 1. The process of hyperparameter-tuning is described in Supplementary Materials and resulted in an initial learning rate of 10^{-5} and batch size of 8. Whilst model training, data augmentation was applied. The images were flipped horizontally, vertically, and horizontally and vertically, in order to obtain a dataset of four times its size. Ten percent of the training data was randomly used for validation.

2.4 Generation of synthetic X-ray CT data

The main problem in training DL models is the lack of a large labeled training dataset. To enlarge the training set and improve segmentation results of pear fruit tissue, synthetic data was generated using a conditional Generative Adversarial Network (cGAN).

In previous work, a DL model to synthesize reliable CT scans from 'Conference' pears was developed (Tempelaere et al., 2023). The model can generate both healthy and defect fruit with different disorder severities. An annotated voxelized 3D drawing of a pear, with labeled regions of 'healthy tissue', 'core', 'internal browning', 'cavities', and 'external air', was given as conditional input to the model, from which the model could generate a reliable CT scan, called a 'synthetic' data entry. More details on this approach are described in Tempelaere et al. (2023).

In the collected CT data of disordered pears, cavity and browning defects always occurred together in the fruit flesh. However, Lammertyn et al. (2003) found that cavities develop from severe internal browning, which means that internal browning could be considered as the first stage of internal quality deterioration. In addition, cavities could also occur without browning of the fruit flesh. That biological variability should thus be included in the dataset. Because of the lack of real pears with only browning or only cavities in our collected CT data (section 2.1), synthetic defect pears were created .

The cGAN was used to extend the available dataset. The conditional nature of the generative model enabled to precisely indicate the location and type of disorder, i.e., browning or cavity. However, it is essential that the these correspond with the existence in nature. In accordance with the finding from Lammertyn et al. (2003), synthetic pear samples were generated that lead from browning or cavities. The workflow started from the CT data available for defect pears that were 'defect and consumable' and 'defect and non-consumable', which all had both browning and cavities, and is illustrated in Fig. 2. For the conditional input of synthetic pears with only browning, the 'cavity' pixels were merged with 'browning' in the original annotated volumes (Fig. 2b). This was assumed to correspond to the first stage of internal browning development in pear fruit (Franck et al., 2007; Lammertyn et al., 2003). For synthetic pears with only cavities, the volume annotations from the 'browning' pixels were changed to ‘healthy tissue’ (Fig. 2c). In total, synthetic CT data of 36 pears with browning and 36 pears with cavities were generated. The conditional semantics were developed in MATLAB (2020a, The MathWorks, Natick, MA, USA) and the generator was implemented in Python using the PyTorch framework version 1.11.0 and ran using a Nvidia Quadro RTX 5000 GPU.

2.5 Experiments

2.5.1 Experiment I: real data

The initial empirical training dataset consisted of 14 healthy pears and 22 pears with both cavities and browning (10 defect and consumable, 12 defect and non-consumable). With 50 radiographs per sample and data augmentation via image flipping, this resulted in a dataset of 7,200 radiographs, from which ten percent was randomly used for validation.

2.5.2 Experiment II: real with additional synthetic data

The prior empirical training dataset was extended with radiography data from the synthetic pears in section 2.4. From the 22 defect pears originally available in the real dataset and consisting of both browning and cavities, a pear with only browning and only cavities was generated. To create a more balanced dataset, additional healthy pears were included for which semantics were predicted in our earlier work (Van De Looverbosch et al., 2021). As the ground-truth labeling was not manually performed, these data was also regarded as additional synthetic data. This resulted in training set of, in total, 36 healthy pears, 22 with both browning and cavities, 22 with only browning, and 22 with only cavities. Radiograph simulation and image flipping resulted in a dataset of 20,400 radiographs, from which ten precent was used for validation.

Prior to experiment I, a model was developed on the same empirical training set without image flipping in the training phase to consider the effect of using data augmentation. In another experiment, the default weighted cross-entropy loss was replaced by IoU loss for possible performance enhancement (Van Beers et al., 2019). More details and results from these two additional experiments are given in Supplementary Materials.

2.6 Model performance

The model performance was evaluated by visual observation and quantitative metrics. The predicted images were thus visually compared with the ground-truth semantics as simulated by the newly introduced method (section 2.2). We focused on the model’s ability to distinguish browning and cavities from the healthy tissue. In addition, pseudo-prediction could occur, which means that during manual annotation of the pear CT data, core and cavities were sometimes difficult to separate and, therefore, labeled as cavities. However, our segmentation model might be able to distinguish them in its predictions.

For the quantitative evaluation, the model predictions were scored by the Intersection over Union score (*IoU*), defined as

$$IoU = \frac{\text{TP}}{\text{FP+TP+FN}} \quad \text{(eq. 1)}$$

with *TP* the number of true positive pixels, *FP* false positives, and *FN* false negatives (Cordts et al., 2016). An *IoU* score equal to 1 indicates a full overlap between the ground-truth and the predicted segmentation.

This metric was calculated per image prediction and for each class, i.e. $IoU_{\text{class i}}$. If the class was not present in both ground-truth and prediction, the calculation of *IoU* would result in zero divided by zero. As the predicted segmentation was correct in this case, an *IoU* score of 1 was assigned to this class. In addition, the mean *IoU* was calculated over the 5 classes per prediction, i.e., $IoU_{\text{radiograph}}$. Statistics were performed to test the statistical significance between the predictions of the different models based on real and additional synthetic data. A repeated measure one-way ANOVA was used for multiple comparisons. The ANOVA test produced *p*-values which were corrected by the Bonferroni correction (in case of equal variances) or the Greenhouse-Geisser correction (in case of unequal variances). A *p*-value of 0.05 was chosen for statistical significance. The level of significance was indicated in the figures as: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

The test set consisted of radiography data for a separate set of real pears, i.e., five healthy, five defect & consumable, and five defect & non-consumable samples and was used for the model based on real data and additional synthetic data. Furthermore, an extra evaluation was performed on pears with only browning or cavities. These samples originated from defect pears, initially consisting of both browning and cavities and not earlier used in the training or validation phase. The samples with only browning or cavities were obtained via the approach in Fig. 2 (section 2.4).

2.7 Benchmark: state-of-the-art classifiers

To evaluate the utility of the segmentation model, the semantic predictions were compared to the heatmaps of two classifiers from our earlier work (Van De Looverbosch et al., 2022). More specifically, the ResNet18 classifier with binary cross-entropy (BCE) loss and the Fully Convolutional Data Description (FCDD) model were used here as a benchmark, as they reached the highest classification success to distinguish the radiographs from healthy and defect pears. Both models were trained in a supervised way. Radiographs of normal pears were simulated from the available CT dataset. The defect samples were obtained by applying outlier exposure on healthy pears, whereby ellipsoidal blobs with a lower grayvalue were added to the normal radiographs. The FCDD model consisted of convolutional and pooling layers, ultimately resulting in a heatmap (Liznerski et al., 2020). This heatmap indicated the features that the model based on to decide whether the pear is healthy or defect. For the ResNet18 classifier, a gradient-based method was applied to visualize the anomalous regions (Selvaraju et al., 2017), as the model does not naturally produce heatmaps. The benchmark classifiers were compared to the segmentation models in terms of utility of their output. To this end, heatmaps of the test set with real and synthetic samples were produced by the ResNet18 classifier and FCDD model.

3. Results

3.1 Performance of the segmentation models

The performance of the segmentation model based on a small empirical manually labeled dataset was compared to a model trained with additional synthetic pears. First, the predictions were visually observed. Figure 4 gives an overview for the three quality classes of pears: 1) healthy, 2) defect but consumable, and 3) defect and non-consumable pears. Ground-truth pictures of the cut-open pears are presented in Fig. 4 (a). The radiographs of the simulated inline X-ray scanner setup are given in Fig. 4 (b), with ground-truth labels in Fig. 4 (c), and the predictions of the models are given in Fig. 4 (d, e). Some general trends can be observed. For the healthy pears, the tissue and core were always accurately predicted by each of the two models, while the defect pears were more difficult to predict. The segmentation of their defect zones was less accurate, in particular for Fig. 4 (d). This model on real data had the tendency to falsely predict browning around the cavity regions. However, extending the training set with synthetic data improved the performance as seen in the results from Fig. 4 (e).

The performance of the models was also quantitatively analyzed in Fig. 5. It presents the IoU for each predicted label for the 50 radiographs of each pear in the test set with real pears, i.e., external air/background, healthy tissue, core, cavity, browning. In addition, the average of those values was calculated for the complete radiograph. The results are grouped for the three different quality classes of pears. Statistical tests for the healthy pears (Fig. 5 (a)) demonstrate that training on additional synthetic data resulted in a higher IoU for the complete radiographs (0.980 ± 0.034), compared to the model only trained on real data (0.957 ± 0.076), resulting from an elevated IoU for cavity and browning. Fig. 5 (b) depicts the results for the pears that were defect but still consumable. Considering the IoU for the complete radiograph, training with additional synthetic data (0.883 ± 0.088) resulted in a significantly better performance than the other model (0.781 ± 0.112). The values for the browning label are very different from each other. The defect pears in the test set always contained both cavities and browning. An IoU score of 0 thus means that the model was not able to segment any browning pixel correctly, when browning was present in the ground-truth labeled radiograph. However, it is possible that the browning was not visible on the ground-truth 2D projection due to the overlay of labels (section 2.2), or that the label is perfectly predicted. These both cases result in an IoU score of 1. As observed from the IoU for browning, the model on real data mostly failed to indicate browning correctly. This in contrast to training on additional synthetic data which was able to accurately predict most of the browning pixels. The results for the defect and non-consumable pears are shown in Fig. 5 (c). No clear statistical significance could be found for the IoU of the radiographs, as the models reached a similar value of .0.727 ± 0.108 and 0.738 ± 0.109. Despite the high value for the $IoU_{radiograph}$, which represents the IoU of all the five labels together, the IoU score for the browning label was rather low (respectively, 0.372 ± 0.225 and 0.418 ± 0.325). Another surprising result is the large variance for the metric IoU_{core}. A deeper inspection of the labeled CT data from which ground-truth labeled radiographs were simulated revealed that the core and cavity labels were sometimes difficult to distinguish in the CT data and, therefore, manually labelled as cavities. Thus, the ground-truth simulated pear radiograph sometimes had no core. However, the two models always predicted a core for all the pears. An IoU score of 0 was thus obtained when no core was labeled in the ground-truth radiographs, as simulated from the labeled CT data (section 2.2).

Besides the test on real radiographs, an extra evaluation was obtained for X-ray radiographs from synthetic pears. These synthetic samples were based on real pears that were not earlier included as either training or validation data. The extra evaluation consisted of defect pears with only browning or cavities. Figure 6 depicts the radiograph, ground-truth labels, and the semantics predicted by the two models. It illustrates that the model on real data was not able to accurately detect brown regions in the radiographs (Fig. 6 (c)). Additional synthetic samples in the training set could, however, drastically improve the model’s detection capability for brown pears (Fig. 6 (d)). The results for the pears with only cavities show that the model on real data falsely predicted a brown region around the cavities (Fig. 6 (c)). The model on additional synthetic data also had this tendency, but the wrongly predicted brown region was smaller (Fig. 6 (d)).

The performance of the models for this synthetic dataset was also quantitively analyzed in Figure 7. The results for the pears with only browning in Fig. 7 (a) and pears with only cavities in Fig. 7 (b) demonstrate a higher performance for the IoU among the complete radiographs when the segmentation model was trained on additional synthetic data. Statistics demonstrated that all labels were better predicted when the segmentation model was trained with additional synthetic data.

Besides the bar plots in Figure 4 and 6, detailed performance values for the two models are provided in Supplementary Materials. In addition, results are provided for an experiment without image flipping and using the IoU loss whilst training instead of the weighted cross-entropy.

3.2 Utility of segmentation maps compared with heatmaps from benchmark classifiers

To the best of the author's knowledge, it is the first time that X-ray radiographs of pear fruit have been segmented by using DL. Prior research mainly focused on binary classification of healthy and defect food (Ahmed et al., 2020; da Silva et al., 2021; Medeiros et al., 2021; Van De Looverbosch et al., 2022; Xue et al., 2022). In this context, the utility of segmentation maps was compared to the output from two state-of-the-art classifiers, i.e., the ResNet18 classifier and FCDD model, developed in our earlier work (Van De Looverbosch et al., 2022). Segmentation and classification tasks are difficult to compare by one-on-one images as for both tasks thresholds should be defined that would influence the model performance. However, our earlier classifiers produced heatmaps that represent regions considered as anomalous and that were assumed to correspond with the browning and cavity regions. To this end, the heatmaps for the radiographs in the test set, as shown in Fig. 8, are visually compared to the segmentation results obtained by the two models in Fig. 4 and 6. For the results of the ResNet18 classifier (Fig. 8 (b)), the pear outline and calyx end are highlighted as anomalous, also in the healthy fruit. For the defect & consumable and defect & non-consumable pears, cavities seem to be indicated as a defect region in addition to the core and pear outline. The heatmaps cannot fully indicate the defects in the pears with only browning, as the heatmaps mainly highlighted the core and pear outline as defect. For the pears with cavities, the heatmap indicated an anomalous region that might correspond to the defect. The heatmaps obtained from the FCDD model are shown in Fig. 8 (c). Here, it could not be decided if the heatmaps can indicate the cavities or browning apart from the core, as the highlighted region was smoothed out around the core. This was always the case for both the healthy and defect pears.

In addition, the heatmaps of the defect & non-consumable pears, which had both browning and cavities, were compared one-on-one with pears with only browning and only cavities, as these latter synthetic samples were produced from the same test samples. The synthetic pears with only browning originated from pears in the test set in which the labels browning and cavities were merged, while the synthetic pears with only cavities originated from pears in the test set in which the label browning was considered as healthy tissue (section 2.4). The comparison of the heatmaps demonstrates that browning could not be visually observed in the heatmaps since the pears with cavities resulted in almost completely the same heatmaps as when the browning was kept in the radiographs.

4. Discussion

4.1 Availability of CT scans significantly improves ground truth segmentation of radiographs

X-ray radiographs and their labelled versions were simulated from pear CT scans and their segmentations, respectively. The simulation approach provided considerable benefits as radiographs from various machine settings, and projection angles can be easily obtained in this manner. For instance, the simulation framework allows to set the source-detector distance, the detector size, and the noise added to the data (van Aarle et al., 2015). In our work, similar settings were chosen as a realistic industrial X-ray scanner and 50 radiographs from different angles were collected per pear (section 2.2).

A new approach was introduced to simulate ground-truth radiographs from the labeled CT scans. For every radiograph, labels could thus be immediately provided. It is a fast and automatic method, but on the other hand, it still requires the CT scans to be (often partially manually) labeled. Therefore, we also state that the model performance is dependent on the accuracy through which the CT scans are labeled. A good example that illustrates this is the pseudo-prediction of the core. To explain, in some CT scans it was difficult to manually distinguish between the core and cavities in the CT data and, therefore, their total volume was labeled as cavity. This was thus also present in the simulated ground-truth semantics obtained from the labeled CT data. However, our model always predicted the core separated from the cavities. This might be considered as a benefit of our model, but needs a more in-depth study of the available pear samples as some fruit might not always have a core either. Another point of attention is the specific order of the overlayed labels in the ground-truth semantics, in which we assumed the core and, secondly, the cavities dominated, followed by the browning and complete pear shape. This can explain that the browning label was sometimes not visible in the ground-truth semantics, while the pear CT data contained browning. In future work, this can be resolved by not limiting the labels to a single channel image, but by holding a binary mask for every semantic entity.

4.2 Synthetic data improves the performance of the segmentation model

Different data augmentation techniques were tested to improve the robustness of our segmentation model. First, image flipping was introduced, which increased the size of the dataset four times. However, this technique only slightly improved the prediction accuracy for defect and non-consumable pear samples, as well was its robustness low (Supplementary Materials). The model was unable to predict the browning zone in pears with browning as their only defect. On the other hand, the model was able to recognize cavities, but it always falsely predicted a browning zone around these cavities. A possible reason was that the training data consisted of healthy pears and defect pears with both browning and cavities. To make the model able to detect cavities and browning independently from each other, the training set was extended with radiographs from pears with only browning and only cavities. These pears were synthesized by our earlier defined generator in Tempelaere et al. (2023) (section 2.4). As conditional input, the labels in the available segmentation maps of the defect pears were merged to obtain pears with only browning and only cavities (Fig. 2).

Using these data as additional training data drastically improved the model performance. Fig. 6 (d) shows that the browning area could be accurately detected. However, the model still predicted some browning around the cavities. Cavities usually develop from severely brown tissue (Franck et al., 2007; Lammertyn et al., 2003). The misclassification of browning around cavities might, hence, be acceptable in practice, as fruit with cavities as their only defect are rather rare. Also, the success of our model depends on the performance of our generator in Tempelaere et al. (2023). As the generator was trained on the same dataset as in this work, which had only defect pears with both cavities and browning, it is possible that the generator is less accurate at synthesizing pears with only one of these two labels. In addition, the synthetic data may represent the same underlying data distribution as already available in our training set. On the other hand, the generator was forced to produce synthetic samples with only one defect due to the semantics as conditional input. As proven in Fig. 5 and 6, these synthetic data could still enhance the model performance.

In future experiments, the generator for synthetic data can indeed be used in combination with another pear dataset to further improve the model performance and robustness. In addition, synthetic pears with only cavities or browning were now obtained by merging labels in the available semantics (Fig. 2). However, the generator might be further optimized by gaining a better understanding of the location, shape, and severity of reliable disorders, as this information is used as conditional input in Tempelaere et al. (2023). For instance, statistical shape models (Danckaers et al., 2017) or a 3D geometrical model generator (Rogge et al., 2015) might be applied to generate new realistic pear shapes, whether or not combined with adding “cavities” and/or “internal browning” to obtain disordered or healthy pears in the end. In addition to synthetic data, other augmentation methods worth considering include cutmix and mixup (Yun et al., 2019; Zhang et al., 2018). Cutmix involves cutting out fragments of one image and placing them over another image, while mixup involves overlaying two images. Finally, the model should not only be tested on synthetic data, but also on real pear samples with only browning or cavities. Unfortunately, those real samples were not available in this study. On the other hand, the higher performance was also visible on the available test set of real pears with browning and cavities. The segmentation model should also be tested on pears from another season and origin, as well on real radiographs instead of simulated radiographs.

4.3 Segmentation maps are highly interpretable results

Our segmentation model was compared in terms of utility to two anomaly detection models, i.e., a ResNet18 classifier and FCDD model (Van De Looverbosch, 2022) as a benchmark. The two benchmark models were trained in a supervised way on data of healthy samples in which anomalies were introduced at random, i.e., in the form of blobs to mimic real defects, using outlier exposure. From the two benchmark models, heatmaps were produced that indicate the regions considered as anomalous. As illustrated in Fig. 8, the cavities were highlighted, but the model failed to recognize the browning. In addition, the pear outline and core were sometimes indicated as defect regions. For the FCDD model, one large region was detected as defect assuming to capture the core, cavities, and browning, but the exact defect region was not clearly defined. In addition, the heatmaps from the healthy pears indicated similar regions as the heatmaps from the defect pears. Comparing the heatmaps to our semantics in Fig. 4 and 6, the segmentation model trained on additional synthetic data seemed to perform better, especially on the pears with only browning and no cavities. This was already suggested in our earlier work (Van De Looverbosch et al., 2022), as the classification success decreased with a smaller cavity volume for pears with both browning and cavities. Possibly, the anomalous samples in the earlier work, which were obtained by outlier exposure, resemble cavities better than browning. The anomalous blobs were ellipsoidal and randomly placed inside the pear, with grayscale values normalized between [0.0, 0.1]. The blobs thus represented a large difference in X-ray attenuation for that specific position. Looking at defects in fruit, cavities also cause a large change in X-ray attenuation compared to browning. Cavities can thus more easily be observed in the radiographs.

We could conclude that the augmentation technique with synthetic data obtained from our generator (Tempelaere et al., 2023) is beneficial as it can introduce more biological variation on defects in our dataset. We argue that these synthetic data better represent the browning defect than the anomalous blobs in case of outlier exposure (Van De Looverbosch et al., 2022).

4.4 Design of the segmentation model compared to other approaches

Earlier work on internal quality of horticultural products by X-ray radiography usually focused on classifying good and bad-quality products (Ahmed et al., 2020; da Silva et al., 2021; Medeiros et al., 2021; Shahin et al., 2002, 2001; Sherif et al., 2023; van Dael et al., 2016; Xue et al., 2022). To the best of our knowledge, it is the first time that, specifically for segmentation, a DL model was developed for radiography data of horticultural products. For each input image, the segmentation model predicts a segmentation map with external air, healthy tissue, core, browning, and cavities as separate labels. The output thus immediately indicates the type, location, and size of the defect. In addition, unlike traditional ML, our DL approach does not require to engineer hand-crafted features to describe the image data. This possibly improves the model performance and speeds up the process of algorithm development compared to earlier studies (Belin et al., 2011; Finney and Norris, 1978; Shahin et al., 2002, 2001; Sherif et al., 2023; van Dael et al., 2016).

In addition, the DL-based segmentation approach is beneficial compared to our benchmark models, i.e., the ResNet18 classifier and FCDD model, whereby the heatmaps indicate defect regions, but do not make a difference between the kind of defects and the core. However, the benchmark models might still provide an indication of the disorder severity by considering the anomaly score. In this case, a threshold value is required as not only the browning and cavities were recognized as anomalous, but the core and pear outline as well. In the earlier work (Van De Looverbosch et al., 2022), this threshold was identified by the Youden index, which corresponds to a threshold value that maximizes the performance for classifying the pears as healthy and defect (Youden, 1950). Apart from the Youden index, the optimal threshold may also depend on the acceptable trade-off between the true positive and false negative rates as specified by the industry. A similar approach could be applied on our predicted segmentation maps. Conditions on the acceptable percentage, location, and shape of browning and/or cavities in the fruit can be defined in practice. However, it is very difficult to define a same threshold for the anomaly heatmaps as for the segmentation maps. Therefore, a one-on-one comparison for image classification based on these two approaches was not performed in this work, but heatmaps and segmentation maps were qualitatively compared.

The current research applies a U-Net architecture, a typical state-of-the-art neural network for semantic segmentation. To date, many variants on this architecture are available, often optimized for a specific application. Some variants, among many others, are nnUNet, ResUNet, and ResUNet++ (Chavan et al., 2022). In addition, it might also be useful to apply multi-task DL to simultaneously produce the segmentation and classification output (Tran et al., 2022). In future research, these models might further improve the network accuracy. On the other hand, other researchers argued that prior studies mainly focus on creating more sophisticated architectures, while significant improvements can already be obtained by adapting another loss function (Janocha and Czarnecki, 2016; Van Beers et al., 2019). Most state-of-the-art neural networks are often still trained on the cross-entropy loss, while the IoU is used as a measure of success in evaluation. For instance, Van Beers et al. (2019) immediately implemented the IoU as loss as introduced by Rahman & Wang (2016) and compared the IoU loss and cross-entropy loss for training a segmentation model on different datasets. Their results showed that training on the IoU loss could significantly increase model performance. We applied the same approach and replaced the cross-entropy loss by the IoU loss. However, no clear improvement in model performance could be observed (Supplementary Materials). A possible explanation is that the IoU loss is a non-decomposable function that consists of several statistics across the training metrics from which gradients are calculated and combined in one loss function. Therefore, direct optimization of the IoU loss can be very difficult compared to traditional losses, such as cross-entropy, that give more direct feedback (L. Yu et al., 2022).

4.5 Industrial relevance of fruit quality inspection

The findings from the current study leads to an overall evaluation of the model performance focused on a practical implementation. Internal quality control is an important aspect in the agri-food industry (Butz et al., 2005). First of all, the industry aims to accurately the detection of defect fruit, leading otherwise to a decreased consumer's appreciation and willingness to pay. In addition, fruit with minor defects could be reprocessed in other food products, reducing food loss. Because of this reason, determining the type, location, and size of defects in fruit is very useful. At the same time, healthy fruit should not be discarded, leading to financial loss and elevated food loss. From a technical point of view, the proposed X-ray technology is cheap, fast, and can be easily implemented in an inline system. X-ray imaging is already used in the food industry and found to be safe to human food products, because of its low voltage and the short exposure time (Mohd Khairi et al., 2018).

Although the segmentation model performs well on simulated data, the model should be further tested on radiography data collected by an industrial X-ray scanner. In addition, the method could be tested on fruit from other cultivars, seasons, and origins. Similar technology and algorithms could be designed for other types of fruit and vegetables, that might lead from internal quality deterioration due to physiological disorders, fungal and bacterial diseases, or pest infection (Snowdon, 2012).

5. Conclusion

The results in this work play a crucial role in advancing automatic internal quality control in the agri-food industry. The paper demonstrated to effectively synthetize X-ray data from defect pear fruit using a conditional Generative Adversarial Network (cGAN). Additionally, the novel simulation method could immediately provide input-target image pairs of radiography data. Both methods contributed to the development of a segmentation model for radiographs from healthy and defect pears with five labels: external air, healthy tissue, core, cavity, and browning. The segmentation approach identified the type of defect, its dimensions, and location, which is more useful than a healthy/bad classification per fruit sample. In future research, the performance could possibly be further increased by using synthetic data with more variance in the pear shape and its defects, or additional real data. Also, the model's transferability should be studied on pears from other cultivars, seasons and origins, and for real radiographs collected by an industrial X-ray line scanner.

6. CrediT authorship contribution statement

Astrid Tempelaere: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Hoang Minh Phan**: Conceptualization, Methodology, Software, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Tim Van De Looverbosch:** Methodology, Data curation. **Pieter Verboven:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Tinne Tuytelaars:** Conceptualization, Writing – review & editing, Project administration, Funding acquisition. **Bart Nicolai:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

7. Acknowledgements

The authors acknowledge the financial support of Fund for Scientific Research (FWO project Astrid Tempelaere 1S23721N) and KU Leuven (project C3191059).

8. References

Ahmed, M.R., Yasmin, J., Park, E., Kim, G., Kim, M.S., Wakholi, C., Mo, C., Cho, B.K., 2020. Classification of watermelon seeds using morphological patterns of x-ray imaging: A comparison of conventional machine learning and deep learning. Sensors 20, 1–15. https://doi.org/10.3390/s20236753

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

Barcelon, E.G., Tojo, S., Watanabe, K., 1999. X-ray Computed Tomography for Internal Quality Evaluation of Peaches. J. Agric. Eng. Res. 323–330.

Belin, E., Rousseau, D., Léchappé, J., Langlois-meurinne, M., Dürr, C., 2011. Rate-distortion tradeoff to optimize high-throughput phenotyping systems . Application to X-ray images of seeds. Comput. Electron. Agric. 77, 188–194. https://doi.org/10.1016/j.compag.2011.05.002

Butz, P., Hofmann, C., Tauscher, B., 2005. Recent Developments in Noninvasive Techniques for Fresh Fruit and Vegetable Internal Quality Analysis. J. Food Sci. 70, 131–141.

Cantre, D., Herremans, E., Verboven, P., Ampofo-asiama, J., Nicolaï, B., 2014. Characterization of the 3-D microstructure of mango (Mangifera indica L . cv . Carabao) during ripening using X-ray computed microtomography. Innov. Food Sci. Emerg. Technol. 24, 28–39. https://doi.org/10.1016/j.ifset.2013.12.008

Chavan, M., Varadarajan, V., Gite, S., Kotecha, K., 2022. Deep Neural Network for Lung Image Segmentation on Chest X-ray. Technologies 10. https://doi.org/10.3390/technologies10050105

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 3213–3223. https://doi.org/10.14797/mdcvj.1087

da Silva, C.B., Silva, A.A.N., Barroso, G., Yamamoto, P.T., Arthur, V., Toledo, C.F.M., Mastrangelo, T. de A., 2021. Convolutional neural networks using enhanced radiographs for real-time detection of sitophilus zeamais in maize grain. Foods 10. https://doi.org/10.3390/foods10040879

Danckaers, F., Huysmans, T., Dael, M. Van, Verboven, P., Nicolaï, B., Sijbers, J., 2017. Building 3D Statistical Shape Models of Horticultural Products. Food Bioprocess Technol. 10, 2100–2112. https://doi.org/10.1007/s11947-017-1979-z

De Schryver, T., Dhaene, J., Dierick, M., Boone, M.N., Janssens, E., Sijbers, J., van Dael, M., Verboven, P., Nicolai, B., Van Hoorebeke, L., 2016. In-Line NDT with X-Ray CT Combining Sample Rotation and Translation. NDT E Int. 84, 89–98. https://doi.org/10.1016/j.ndteint.2016.09.001

Finney, E.E., Norris, K.H., 1978. X-ray scans for detecting hollow heart in potatoes. Am. Potato J. 95–105.

Franck, C., Lammertyn, J., Ho, Q.T., Verboven, P., Verlinden, B., Nicolaï, B.M., 2007. Browning Disorders in Pear Fruit. Postharvest Biol. Technol. 43, 1–13. https://doi.org/10.1016/j.postharvbio.2006.08.008

Gadgile, D., Chavan, A., 2017. Detection of post-harvest fungal diseases of mango by X-ray scanning non-destructive technology 7, 65–69. https://doi.org/10.5943/ppq/7/1/8

Gadgile, D., Joshi, C.P., Shinde, V.M., Kachare, P.B., 2017. Detection of green mold rot infection of citrus fruit by X-ray scanning non-destructive technology 8, 78–80. https://doi.org/10.19071/cb.2017.v8.3211

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning, 1st ed. Mit Press Ltd.

Herremans, E., Melado-Herreros, A., Defraeye, T., Verlinden, B., Hertog, M., Verboven, P., Val, J., Fernández-Valle, M.E., Bongaers, E., Estrade, P., Wevers, M., Barreiro, P., Nicolaï, B.M., 2014. Comparison of X-ray CT and MRI of Watercore Disorder of Different Apple Cultivars. Postharvest Biol. Technol. 87, 42–50. https://doi.org/10.1016/j.postharvbio.2013.08.008

Janocha, K., Czarnecki, W.M., 2016. On loss functions for deep neural networks in classification. Schedae Informaticae 25, 49–59. https://doi.org/10.4467/20838476SI.16.004.6185

Janssens, E., Alves Pereira, L.F., De Beenhouwer, J., Tsang, I.R., Van Dael, M., Verboven, P., Nicolaï, 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

Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization. 3rd Int. Conf. Learn. Represent. ICLR 2015 1–15.

Lammertyn, J., Dresselaers, T., Van Hecke, P., Jancsók, P., Wevers, M., Nicolaï, B.M., 2003. 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

Liznerski, P., Ruff, L., Vandermeulen, R.A., Franks, B.J., Kloft, M., Müller, K.-R., 2020. Explainable Deep One-Class Classification 1–25.

Medeiros, A.D. De, Cupertino, R., Barboza, C., 2021. Deep learning-based approach using X-ray images for classifying Crambe abyssinica seed quality. Ind. Crop. Prod. 164. https://doi.org/10.1016/j.indcrop.2021.113378

Medeiros, A.D. De, Pinheiro, D.T., Xavier, W.A., Junio, L., Cunha, D., Dias, S., 2020. Quality classification of Jatropha curcas seeds using radiographic images and machine learning. Ind. Crop. Prod. 146, 112162. https://doi.org/10.1016/j.indcrop.2020.112162

Mohd Khairi, M.T., Ibrahim, S., Md Yunus, M.A., Faramarzi, M., 2018. Noninvasive techniques for detection of foreign bodies in food: A review. J. Food Process Eng. 41. https://doi.org/10.1111/jfpe.12808

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. 5, 285–312. https://doi.org/10.1146/annurev-food-030713-092410

Ogawa, Y., Kondo, N., Shibusawa, S., 2003. Inside Quality Evaluation of Fruit by X-ray Image. Proc. 2003 IEEUASME Int. Conf. Adv. Intell. Mechatron. 1360–1365.

Otsu, N., Smith, P.L., Reid, D.B., Environment, C., Palo, L., Alto, P., Smith, P.L., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man. Cybern. C, 62–66.

Poly, K., Wako, O., Fumina, T., Fumihiko, T., 2018. Quality Evaluation of 1-Methylcyclopropene Treated Japanese Apricot Using X-ray Computed Tomography. Food Sci. Technol. Res. 24, 377–384. https://doi.org/10.3136/fstr.24.377

Rady, A.M., Guyer, D.E., Donis-González, I.R., Kirk, W., Watson, N.J., 2020. A comparison of different optical instruments and machine learning techniques to identify sprouting activity in potatoes during storage. J. Food Meas. Charact. https://doi.org/10.1007/s11694-020-00590-2

Rahman, M.A., Wang, Y., 2016. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10072 LNCS, 234–244. https://doi.org/10.1007/978-3-319-50835-1_22

Rogge, S., Defraeye, T., Herremans, E., Verboven, P., Nicolaï, B.M., 2015. A 3D Contour Based Geometrical Model Generator for Complex-Shaped Horticultural Products. J. Food Eng. 157, 24–32. https://doi.org/10.1016/j.jfoodeng.2015.02.006

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Schatzki, T.F.;, Haff, R.P.;, Young, R.;, Can, I.;, Le, L.C.;, Toyofuku, N., 1996. Defect detection in apples by means of x-ray imaging. Proc. SPIE 2907, 176–185.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., Localization, G., Dec, C. V, Cogswell, M., 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Proc. IEEE Int. Conf. Comput. Vis. 218, 336–359.

Shahin, M.A., Tollner, E.W., McClendon, R.W., 2001. Artificial Intelligence Classifiers for Sorting Apples based on Watercore. J. Agric. Eng. Res. 79, 265–274. https://doi.org/10.1006/jaer.2001.0705

Shahin, M.A., Tollner, E.W., Mcclendon, R.W., Arabnia, H.R., 2002. APPLE CLASSIFICATION BASED ON SURFACE BRUISES USING IMAGE PROCESSING AND NEURAL NETWORKS. Am. Soc. Agric. Eng. 45, 1619–1627.

Sherif, H., Charrier, A., Rasti, P., Guiloteau, E., Messaka, J., Rousseau, D., 2023. Automatic fasciation detection in salad with 2D X-ray imaging. Acta Hortic. 225–228. https://doi.org/10.17660/actahortic.2023.1360.28

Snowdon, A.L., 2012. A Colour Atlas of Post-Harvest Diseases and Disorders of Fruits and Vegetables: Vegetables v. 2. Manson Publishing Ltd.

Tempelaere, A., Van De Looverbosch, T., Kelchtermans, K., Verboven, P., Tuytelaars, T., Nicolai, B., 2023. Synthetic data for X-ray CT of healthy and disordered pear fruit using deep learning. Postharvest Biol. Technol. 200, 112342. https://doi.org/10.1016/j.postharvbio.2023.112342

Tran, V.L., Lin, H.Y., Liu, H.W., 2022. Multitask Deep Learning for Segmentation and Lumbosacral Spine Inspection. IEEE Trans. Instrum. Meas. 71, 1–10. https://doi.org/10.1109/TIM.2022.3184341

van Aarle, W., Palenstijn, W.J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K.J., Sijbers, J., 2015. The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography. Ultramicroscopy 157, 35–47. https://doi.org/10.1016/j.ultramic.2015.05.002

Van Beers, F., Lindström, A., Okafor, E., Wiering, M.A., 2019. Deep neural networks with intersection over union loss for binary image segmentation. ICPRAM 2019 - Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods 438–445. https://doi.org/10.5220/0007347504380445

van Dael, M., Lebotsa, S., Herremans, E., Verboven, P., Sijbers, J., Opara, U.L., Cronje, P.J., Nicolaï, B.M., 2016. A Segmentation and Classification Algorithm for Online Detection of Internal Disorders in Citrus Using X-Ray Radiographs. Postharvest Biol. Technol. 112, 205–214.

https://doi.org/10.1016/j.postharvbio.2015.09.020

Van De Looverbosch, T., He, J., Tempelaere, A., Kelchtermans, K., Verboven, P., Tuytelaars, T., Sijbers, J., Nicolai, B., 2022. Inline nondestructive internal disorder detection in pear fruit using explainable deep anomaly detection on X-ray images. Comput. Electron. Agric. 197, 106962. https://doi.org/10.1016/j.compag.2022.106962

Van De Looverbosch, T., Raeymaekers, E., Verboven, P., Sijbers, J., Nicolaï, B., 2021. Non-destructive internal disorder detection of Conference pears by semantic segmentation of X-ray CT scans using deep learning. Expert Syst. Appl. 176, 114925. https://doi.org/10.1016/j.eswa.2021.114925

Van De Looverbosch, T., Rahman Bhuiyan, M.H., Verboven, P., Dierick, M., Van Loo, D., De Beenbouwer, J., Sijbers, J., Nicolaï, B., 2020. Nondestructive Internal Quality Inspection of Pear Fruit by X-ray CT using Machine Learning. Food Control 113, 107170. https://doi.org/10.1016/j.foodcont.2020.107170

Watkins, C., Olatunji, J.R., East, A.R., DeLong, J., Prange, R., Brandenburg, J., More, and 31, 2020. Advances in postharvest management of horticultural produce.

Xue, Q., Miao, P., Yu, Y., 2022. X-ray-based machine vision technique for detection of internal defects of sterculia seeds. J. Food Sci. 3386–3395. https://doi.org/10.1111/1750-3841.16237

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3, 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Yu, L., Li, Z., Xu, M., Gao, Y., Luo, J., Zhang, J., 2022. Distribution-Aware Margin Calibration for Semantic Segmentation in Images. Int. J. Comput. Vis. 130, 95–110. https://doi.org/10.1007/s11263-021-01533-0

Yu, S., Wang, N., Ding, X., Qi, Z., Hu, N., Duan, S., Yang, Z., Bi, X., 2022. Detection of pear freezing injury by non-destructive X-ray scanning technology. Postharvest Biol. Technol. 190, 111950. https://doi.org/10.1016/j.postharvbio.2022.111950

Yun, S., Han, D., Chun, S., Oh, S.J., Choe, J., Yoo, Y., 2019. CutMix: Regularization strategy to train strong classifiers with localizable features. Proc. IEEE Int. Conf. Comput. Vis. 2019-Octob, 6022–6031. https://doi.org/10.1109/ICCV.2019.00612

Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D., 2018. MixUp: Beyond empirical risk minimization. 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc. 1–13.

Zhang, Z., Sabuncu, M.R., 2018. Generalized cross entropy loss for training deep neural networks with noisy labels. Adv. Neural Inf. Process. Syst. 2018-Decem, 8778–8788.

(Zhang and Sabuncu, 2018)

Ahmed, M.R., Yasmin, J., Park, E., Kim, G., Kim, M.S., Wakholi, C., Mo, C., Cho, B.K., 2020. Classification of watermelon seeds using morphological patterns of x-ray imaging: A comparison of conventional machine learning and deep learning. Sensors 20, 1–15. https://doi.org/10.3390/s20236753

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

Barcelon, E.G., Tojo, S., Watanabe, K., 1999. X-ray Computed Tomography for Internal Quality Evaluation of Peaches. J. Agric. Eng. Res. 323–330.

Belin, E., Rousseau, D., Léchappé, J., Langlois-meurinne, M., Dürr, C., 2011. Rate-distortion tradeoff to optimize high-throughput phenotyping systems . Application to X-ray images of seeds. Comput. Electron. Agric. 77, 188–194. https://doi.org/10.1016/j.compag.2011.05.002

Butz, P., Hofmann, C., Tauscher, B., 2005. Recent Developments in Noninvasive Techniques for Fresh Fruit and Vegetable Internal Quality Analysis. J. Food Sci. 70, 131–141.

Cantre, D., Herremans, E., Verboven, P., Ampofo-asiama, J., Nicolaï, B., 2014. Characterization of the 3-D microstructure of mango (Mangifera indica L . cv . Carabao) during ripening using X-ray computed microtomography. Innov. Food Sci. Emerg. Technol. 24, 28–39. https://doi.org/10.1016/j.ifset.2013.12.008

Chavan, M., Varadarajan, V., Gite, S., Kotecha, K., 2022. Deep Neural Network for Lung Image Segmentation on Chest X-ray. Technologies 10. https://doi.org/10.3390/technologies10050105

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 3213–3223. https://doi.org/10.14797/mdcvj.1087

da Silva, C.B., Silva, A.A.N., Barroso, G., Yamamoto, P.T., Arthur, V., Toledo, C.F.M., Mastrangelo, T. de A., 2021. Convolutional neural networks using enhanced radiographs for real-time detection of sitophilus zeamais in maize grain. Foods 10. https://doi.org/10.3390/foods10040879

Danckaers, F., Huysmans, T., Dael, M. Van, Verboven, P., Nicolaï, B., Sijbers, J., 2017. Building 3D Statistical Shape Models of Horticultural Products. Food Bioprocess Technol. 10, 2100–2112. https://doi.org/10.1007/s11947-017-1979-z

De Schryver, T., Dhaene, J., Dierick, M., Boone, M.N., Janssens, E., Sijbers, J., van Dael, M., Verboven, P., Nicolai, B., Van Hoorebeke, L., 2016. In-Line NDT with X-Ray CT Combining Sample Rotation and Translation. NDT E Int. 84, 89–98. https://doi.org/10.1016/j.ndteint.2016.09.001

Finney, E.E., Norris, K.H., 1978. X-ray scans for detecting hollow heart in potatoes. Am. Potato J. 95–105.

Franck, C., Lammertyn, J., Ho, Q.T., Verboven, P., Verlinden, B., Nicolaï, B.M., 2007. Browning Disorders in Pear Fruit. Postharvest Biol. Technol. 43, 1–13. https://doi.org/10.1016/j.postharvbio.2006.08.008

Gadgile, D., Chavan, A., 2017. Detection of post-harvest fungal diseases of mango by X-ray scanning non-destructive technology 7, 65–69. https://doi.org/10.5943/ppq/7/1/8

Gadgile, D., Joshi, C.P., Shinde, V.M., Kachare, P.B., 2017. Detection of green mold rot infection of citrus fruit by X-ray scanning non-destructive technology 8, 78–80. https://doi.org/10.19071/cb.2017.v8.3211

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning, 1st ed. Mit Press Ltd.

Herremans, E., Melado-Herreros, A., Defraeye, T., Verlinden, B., Hertog, M., Verboven, P., Val, J., Fernández-Valle, M.E., Bongaers, E., Estrade, P., Wevers, M., Barreiro, P., Nicolaï, B.M., 2014. Comparison of X-ray CT and MRI of Watercore Disorder of Different Apple Cultivars. Postharvest Biol. Technol. 87, 42–50. https://doi.org/10.1016/j.postharvbio.2013.08.008

Janocha, K., Czarnecki, W.M., 2016. On loss functions for deep neural networks in classification. Schedae Informaticae 25, 49–59. https://doi.org/10.4467/20838476SI.16.004.6185

Janssens, E., Alves Pereira, L.F., De Beenhouwer, J., Tsang, I.R., Van Dael, M., Verboven, P., Nicolaï, 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

Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization. 3rd Int. Conf. Learn. Represent. ICLR 2015 1–15.

Lammertyn, J., Dresselaers, T., Van Hecke, P., Jancsók, P., Wevers, M., Nicolaï, B.M., 2003. 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

Liznerski, P., Ruff, L., Vandermeulen, R.A., Franks, B.J., Kloft, M., Müller, K.-R., 2020. Explainable Deep One-Class Classification 1–25.

Medeiros, A.D. De, Cupertino, R., Barboza, C., 2021. Deep learning-based approach using X-ray images for classifying Crambe abyssinica seed quality. Ind. Crop. Prod. 164. https://doi.org/10.1016/j.indcrop.2021.113378

Medeiros, A.D. De, Pinheiro, D.T., Xavier, W.A., Junio, L., Cunha, D., Dias, S., 2020. Quality classification of Jatropha curcas seeds using radiographic images and machine learning. Ind. Crop. Prod. 146, 112162. https://doi.org/10.1016/j.indcrop.2020.112162

Mohd Khairi, M.T., Ibrahim, S., Md Yunus, M.A., Faramarzi, M., 2018. Noninvasive techniques for detection of foreign bodies in food: A review. J. Food Process Eng. 41. https://doi.org/10.1111/jfpe.12808

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. 5, 285–312. https://doi.org/10.1146/annurev-food-030713-092410

Ogawa, Y., Kondo, N., Shibusawa, S., 2003. Inside Quality Evaluation of Fruit by X-ray Image. Proc. 2003 IEEUASME Int. Conf. Adv. Intell. Mechatron. 1360–1365.

Otsu, N., Smith, P.L., Reid, D.B., Environment, C., Palo, L., Alto, P., Smith, P.L., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst. Man. Cybern. C, 62–66.

Poly, K., Wako, O., Fumina, T., Fumihiko, T., 2018. Quality Evaluation of 1-Methylcyclopropene Treated Japanese Apricot Using X-ray Computed Tomography. Food Sci. Technol. Res. 24, 377–384. https://doi.org/10.3136/fstr.24.377

Rady, A.M., Guyer, D.E., Donis-González, I.R., Kirk, W., Watson, N.J., 2020. A comparison of different optical instruments and machine learning techniques to identify sprouting activity in potatoes during storage. J. Food Meas. Charact. https://doi.org/10.1007/s11694-020-00590-2

Rahman, M.A., Wang, Y., 2016. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10072 LNCS, 234–244. https://doi.org/10.1007/978-3-319-50835-1_22

Rogge, S., Defraeye, T., Herremans, E., Verboven, P., Nicolaï, B.M., 2015. A 3D Contour Based Geometrical Model Generator for Complex-Shaped Horticultural Products. J. Food Eng. 157, 24–32. https://doi.org/10.1016/j.jfoodeng.2015.02.006

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Schatzki, T.F.;, Haff, R.P.;, Young, R.;, Can, I.;, Le, L.C.;, Toyofuku, N., 1996. Defect detection in apples by means of x-ray imaging. Proc. SPIE 2907, 176–185.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., Localization, G., Dec, C. V, Cogswell, M., 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Proc. IEEE Int. Conf. Comput. Vis. 218, 336–359.

Shahin, M.A., Tollner, E.W., McClendon, R.W., 2001. Artificial Intelligence Classifiers for Sorting Apples based on Watercore. J. Agric. Eng. Res. 79, 265–274. https://doi.org/10.1006/jaer.2001.0705

Shahin, M.A., Tollner, E.W., Mcclendon, R.W., Arabnia, H.R., 2002. APPLE CLASSIFICATION BASED ON SURFACE BRUISES USING IMAGE PROCESSING AND NEURAL NETWORKS. Am. Soc. Agric. Eng. 45, 1619–1627.

Sherif, H., Charrier, A., Rasti, P., Guiloteau, E., Messaka, J., Rousseau, D., 2023. Automatic fasciation detection in salad with 2D X-ray imaging. Acta Hortic. 225–228. https://doi.org/10.17660/actahortic.2023.1360.28

Snowdon, A.L., 2012. A Colour Atlas of Post-Harvest Diseases and Disorders of Fruits and Vegetables: Vegetables v. 2. Manson Publishing Ltd.

Tempelaere, A., Van De Looverbosch, T., Kelchtermans, K., Verboven, P., Tuytelaars, T., Nicolai, B., 2023. Synthetic data for X-ray CT of healthy and disordered pear fruit using deep learning. Postharvest Biol. Technol. 200, 112342. https://doi.org/10.1016/j.postharvbio.2023.112342

Tran, V.L., Lin, H.Y., Liu, H.W., 2022. Multitask Deep Learning for Segmentation and Lumbosacral Spine Inspection. IEEE Trans. Instrum. Meas. 71, 1–10. https://doi.org/10.1109/TIM.2022.3184341

van Aarle, W., Palenstijn, W.J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K.J., Sijbers, J., 2015. The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography. Ultramicroscopy 157, 35–47. https://doi.org/10.1016/j.ultramic.2015.05.002

Van Beers, F., Lindström, A., Okafor, E., Wiering, M.A., 2019. Deep neural networks with intersection over union loss for binary image segmentation. ICPRAM 2019 - Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods 438–445. https://doi.org/10.5220/0007347504380445

van Dael, M., Lebotsa, S., Herremans, E., Verboven, P., Sijbers, J., Opara, U.L., Cronje, P.J., Nicolaï, B.M., 2016. A Segmentation and Classification Algorithm for Online Detection of Internal Disorders in Citrus Using X-Ray Radiographs. Postharvest Biol. Technol. 112, 205–214. https://doi.org/10.1016/j.postharvbio.2015.09.020

Van De Looverbosch, T., He, J., Tempelaere, A., Kelchtermans, K., Verboven, P., Tuytelaars, T., Sijbers, J., Nicolai, B., 2022. Inline nondestructive internal disorder detection in pear fruit using explainable deep anomaly detection on X-ray images. Comput. Electron. Agric. 197, 106962. https://doi.org/10.1016/j.compag.2022.106962

Van De Looverbosch, T., Raeymaekers, E., Verboven, P., Sijbers, J., Nicolaï, B., 2021. Non-destructive internal disorder detection of Conference pears by semantic segmentation of X-ray CT scans using deep learning. Expert Syst. Appl. 176, 114925. https://doi.org/10.1016/j.eswa.2021.114925

Van De Looverbosch, T., Rahman Bhuiyan, M.H., Verboven, P., Dierick, M., Van Loo, D., De Beenbouwer, J., Sijbers, J., Nicolaï, B., 2020. Nondestructive Internal Quality Inspection of Pear Fruit by X-ray CT using Machine Learning. Food Control 113, 107170. https://doi.org/10.1016/j.foodcont.2020.107170

Watkins, C., Olatunji, J.R., East, A.R., DeLong, J., Prange, R., Brandenburg, J., More, and 31, 2020. Advances in postharvest management of horticultural produce.

Xue, Q., Miao, P., Yu, Y., 2022. X-ray-based machine vision technique for detection of internal defects of sterculia seeds. J. Food Sci. 3386–3395. https://doi.org/10.1111/1750-3841.16237

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3, 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Yu, L., Li, Z., Xu, M., Gao, Y., Luo, J., Zhang, J., 2022. Distribution-Aware Margin Calibration for Semantic Segmentation in Images. Int. J. Comput. Vis. 130, 95–110. https://doi.org/10.1007/s11263-021-01533-0

Yu, S., Wang, N., Ding, X., Qi, Z., Hu, N., Duan, S., Yang, Z., Bi, X., 2022. Detection of pear freezing injury by non-destructive X-ray scanning technology. Postharvest Biol. Technol. 190, 111950. https://doi.org/10.1016/j.postharvbio.2022.111950

Yun, S., Han, D., Chun, S., Oh, S.J., Choe, J., Yoo, Y., 2019. CutMix: Regularization strategy to train strong classifiers with localizable features. Proc. IEEE Int. Conf. Comput. Vis. 2019-Octob, 6022–6031. https://doi.org/10.1109/ICCV.2019.00612

Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D., 2018. MixUp: Beyond empirical risk minimization. 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc. 1–13.

Zhang, Z., Sabuncu, M.R., 2018. Generalized cross entropy loss for training deep neural networks with noisy labels. Adv. Neural Inf. Process. Syst. 2018-Decem, 8778–8788.