



## ARCHIVIO ISTITUZIONALE DELLA RICERCA

### Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

A convolutional neural network approach to detecting fruit physiological disorders and maturity in 'Abbé Fétel' pears

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

A convolutional neural network approach to detecting fruit physiological disorders and maturity in 'Abbé Fétel' pears / Bonora A.; Bortolotti G.; Bresilla K.; Grappadelli L.C.; Manfrini L.. - In: BIOSYSTEMS ENGINEERING. - ISSN 1537-5110. - ELETTRONICO. - 212:(2021), pp. 264-272. [10.1016/j.biosystemseng.2021.10.009]

This version is available at: <https://hdl.handle.net/11585/849221> since: 2022-01-31

*Published:*

DOI: <http://doi.org/10.1016/j.biosystemseng.2021.10.009>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

(Article begins on next page)

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

# A Convolutional Neural Network Approach to Detecting Fruit Physiological disorders and Maturity in ‘Abbé Fétel’ pears

Alessandro Bonora,<sup>1</sup> Gianmarco Bortolotti,<sup>1</sup> Kushtrim Bresilla,<sup>1</sup> Luca Corelli Grappadelli<sup>1</sup> and Luigi Manfrini<sup>1</sup>

<sup>1</sup> Department of Agricultural and Food Science, University of Bologna, Bologna, 40127, Italy

Correspondence should be addressed to Alessandro Bonora: [a.bonora@unibo.it](mailto:a.bonora@unibo.it)

## ABSTRACT

Image processing through the implementation of manually coded algorithms has been adopted to detect fruit damage during post-harvest operations. This study tested convolution neural networks with “You Only Look Once” (YOLO) architecture using a commercial online platform to detect physiological disorders in ‘Abbé Fétel’ pear. Disorders such as superficial scald and the related ripening stage using the starch pattern index (SPI) were assessed. Three different models were trained to detect: i) individual fruit within the boxes; ii) superficial scald or senescence scald on pear skin; iii) the SPI value of pears was assessed using the Lugol solution. Preliminary statistics show that the model to count the fruit inside the boxes reached 64.70 % of true positives with 0.5 intersection over union. The second had less accuracy (up to 20 % of true positives) but maintained a good average precision (60 %) with different confidence thresholds (40 % and 20 %). The third showed good performances compared to the Ctifl and Laimburg scales, with an F1 score of 0.36 and 0.59, respectively. The effectiveness of the transfer learning method was demonstrated but further image labelling and modelling research is needed to improve the accuracy of the simulations and to develop an application for portable devices for pre- and post-harvest factor mapping. These results could lead to improvements in the management of fruit boxes and thus help ensure good fruit quality for consumers.

Keywords: *Pyrus communis*, fruit quality, superficial scald, starch pattern index, neural networks, YOLO

## 1. INTRODUCTION

Various technologies have been used to evaluate fruit quality attributes effectively and economically. Although a demand for automation together with better and faster solutions in the post-harvest industry is increasing, only a few of them have been adopted (Bazame, Molin, Althoff, & Martello, 2021; Jayas & Karunakaran, 2005). Computer vision

29 techniques are a non-destructive, low operational cost, reliable, objective and efficient approach to this end, based on  
30 image processing (Leiva, Mondragon, Mery, & Aguilera, 2011). Artificial neural networks (ANN) have proven to be  
31 very effective to identify and classify fruit based on quality, where non-coherence or non-linearity often exists (El Masry  
32 et al., 2009). Moreover, the use of neural networks can help to add value to horticultural management, promoting  
33 productivity, better quality control processes, and achieve flexible farm management (Juan P. Vasconez, Kantor, & Auat  
34 Cheein, 2019). Strategies that allow data acquisition and analysis from agricultural products can help to optimize current  
35 practices, such as yield estimation, pathogen and disease detection and crop maturity classification (J. P. Vasconez,  
36 Delpiano, Vougioukas, & Auat Cheein, 2020). Barreiro et al., (1997) tested different neural network approaches for  
37 bruise prediction in apple, pear and peach and Kim et al., (2000) used non-linear techniques (ANN) based on multilayer  
38 perception with variations on back-propagation learning to classify kiwi fruit stored under different conditions. Kavdir  
39 & Guyer, (2004) used a different approach and developed a back-propagation neural network (BPNN) with the textural  
40 features extracted from spatial distribution of colour (grey colour) levels to detect defects in Empire and Golden  
41 Delicious apples. ANN methods were also used by Gama et al., (2017) to simulate the enzymatic hydrolysis process  
42 with the optimal conditions being successfully selected, as indicated by the  $R^2$  value of 0.99 and a small MSE value.  
43 Superficial scald (SS) is one of the major challenges in the pear sector and it severely affects Abbé Fétel pear during  
44 storage and reduces shelf life (Lurie & Watkins, 2012). Thus, an accurate and rapid detection method for this disease  
45 could help to develop an early treatment technique while substantially reducing economic losses (Fuentes, Yoon, Kim,  
46 & Park, 2017). The severity of SS is related to fruit maturity determined by SPI (Blanpied et al., 1991; Brouwer,  
47 Mensink, Hogeveen-van Echtelt, & Woltering, 2021). A model for SS was developed by Kupferman, (2001) using  
48 harvest date, starch score, and the number of preharvest days below 10 °C. In ‘Nanguo’ pears, n-butanol controlled core  
49 browning and preserved starch grains in the core tissue during storage were observed by Sun et al., (2020) using  
50 transmission electron microscopy. However, SPI is not always employed in pears although some studies have reported  
51 the use of this procedure to estimate fruit storability (Agar, Biasi, & Mitcham, 1999). While some researchers have  
52 observed that pears showed a similar behaviour to apples (Isidoro & Almeida, 2006), but in ‘Abbé Fétel’ and ‘Beurré  
53 D’Anjou’ pears starch degradation is related to the advance of fruit maturity and the superficial scald development  
54 (Calvo et al., 2011; Calvo et al., 2015). Moreover, ‘Bartlett’ pears have reported higher scald incidence in late harvested  
55 fruit (Bower, Biasi, & Mitcham, 2003; Whitaker, Villalobos-Acuña, Mitcham, & Mattheis, 2009). More recently,  
56 Zanella et al., (2019), in collaboration with Isolcell SPA, Laives, BZ, Italy developed an automated starch-meter  
57 (Amilon) that still uses iodine-potassium but integrates an imaging system that quantifies the colorimetric reaction for

58 unbiased results. Further studies are therefore needed to provide a fast, simple, and reliable software for portable devices  
59 to detect disorders and replace completely the traditional starch assessment method for pear fruit during storage. This  
60 work was aimed at training and testing neural network using images to detect superficial scald on individual fruit and  
61 to classify equatorial slices of fresh cut fruit dipped in an iodine solution.

## 62 2. MATERIAL AND METHOD

### 63 2.1. Model Architecture

64 In the study a convolutional neural network (CNN), a specialized type of artificial neural network used for image  
65 analysis, has been adopted. Between different CNNs available, the object detection system called “You Only Look  
66 Once”, and in particular its third version (YOLO v3), was selected because it is a fast and efficient convolutional network  
67 for detection and localisation and is suitable for application in portable devices. Compared to other current methods that  
68 treat detection, classification and region extraction as different problems, YOLO v3 has the advantage of a single step  
69 approach (Redmon & Farhadi, 2018). To achieve that, YOLO v3 sacrifices accuracy to gain speed. It takes as input an  
70 image of max size  $608 \times 608$  pixels and divides it into grid cells (with a dimension  $S \times S$ ). Each grid cell is responsible  
71 for the bounding box whose centre is at the location of the grid cell and predicts bounding boxes (B) as well as  
72 confidence level and class probability. In a dataset with class labels (C), the output tensor is  $S \times S \times (C + B \times 5)$ .  
73 However, the main problem with the standard YOLO v3 is that the model can detect only one object class per cell  
74 (Redmon, Divvala, Girshick, & Farhadi, 2016).

### 75 2.2. Dataset

76 Images were collected in the laboratories of the Department of Agricultural and Food Sciences, University of  
77 Bologna, Italy, and in a storage warehouse in Altedo, Bologna, Italy. From 2018 to 2020, multiple images of Abbé Fétel  
78 pears, sourced from 30 orchards located in the Emilia Romagna Region, Italy were taken at different angles and light  
79 conditions, both with a camera and a smartphone. The orchards were selected to represent a selection of different  
80 features such as age, vigour, rootstock and training system. Concerning pear counting, 122 images of pear boxes were  
81 taken in 2019, and 3,133 fruit were labelled to distinguish them inside a box (number of fruit as independent variable).  
82 On the other hand, 851 images of pear boxes were taken, and 3071 superficial scald symptoms were labelled to detect  
83 the storage disorder from different producers and at different times during the storage seasons 2018-2019 and 2019-  
84 2020 (number of symptoms as independent variable). Finally, 318 pictures of pears were captured, and 4043 fruit with  
85 Lugol solution was used to recognise the different classes of iodine reaction having a well-represented SPI range. The

86 latter dataset was created with images of pear fruit slices taken at the harvest during three consecutive seasons (2018,  
87 2019 and 2020).

### 88 2.3. GPU and Platform

89 To label the images, and train our models, a computer vision platform was used (<https://supervise.ly>). A workstation  
90 equipped with an NVIDIA RTX2080 SUPER- RAM: 8GB card as GPU was used as the computer cluster to train the  
91 algorithm. The dataset was augmented with a DTL (Data Transformation Language) tool, available on the web platform,  
92 to enlarge the data set which helped to reduce over-fitting in supervised learning algorithms (Table 1; Krizhevsky &  
93 Hinton, 2012). The final datasets were divided into training, validation, and test set which was excluded from the training  
94 session. For the SPI model, the test set was a totally independent data set from 2018 and it was extracted before  
95 augmentation. On the contrary, the test pictures to evaluate the other models were randomly chosen as 10 % of the total  
96 augmented pictures and extracted before the training session. Regarding the validation set while training, the algorithm  
97 considered 15% of the remaining photos.

	Training set	Validation set	Test set
Pear counting	4502	804	550
Superficial scald	8466	1446	300
Starch pattern index	3108	348	30

98  
99

Table 1. Number of augmented images used for training and testing the pear fruit ANNs.

100 YOLO v3 (COCO pre-trained version) was operated up to 60 epochs to train the pear counting model (Table 2). The  
101 superficial scald model algorithm was trained for 200 epochs (i.e., two sessions of 100 epochs; Table 2). The starch test  
102 model was instead trained in differently, as shown in Table 2. Firstly, a transfer learning (TL) approach was considered:  
103 the model was improved by adding pictures whilst labelling was underway (123 epochs, 130 epochs and 100 epochs).  
104 Secondly, all the labelled images and all the datasets were trained in a massive training (MT) session with 353 epochs  
105 in total. This procedure was repeated with the starch-iodine dataset to take into account ten classes (Le code Amidon  
106 Pomme, 2002 from Ctifl, France) with one of the five classes coming from the Laimburg Experiment Station (Werth,  
107 2000). For each task, the best models or checkpoints were selected looking at the train graph and the loss value (Fig. 1).

108

	Pear counting	Superficial scald	SPI 1-10 (transfer learning)	SPI 1-10 (massive training)	SPI 1-5 (transfer learning)	SPI 1-5 (massive training)
Training sessions	1	2	3	1	2	1
Epochs	60	200	353	353	353	353

Table 2. Number of training sessions and epochs used for pear fruit ANNs.

To validate our model, 10% of the original datasets with the labelled images were extracted and included in a validation dataset (defined “test set”) that was tested with the trained algorithm (Jarvinen, Choi, Heinemann, & Baugher, 2018). Afterwards, their accuracy was estimated by calculating confusion matrices on the results with imposed intersection over union (IoU). Precision (1), recall (2) and F1 score (3) were calculated as important parameters for fruit detection where TP is the number of true positives (matches), FP is the number of false positives (false detection), and FN is the number of false negatives (missed; Tu et al., 2020). In the case of superficial scald, the test set was validated with two levels of confidence thresholds (0.4 and 0.2) that represent the class probability of the bounding boxes YOLO v3 trained model (Bresilla et al., 2019).

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \quad (3)$$

### 3. RESULTS AND DISCUSSION

#### 3.1. Pear detection model

The output analysis was evaluated using the data of labelled images from the validation dataset. As a first step the model was trained to count pear fruit inside boxes. The matrix evaluated the accuracy for the first model, according to the well-known criteria based on Pascal Visual Object Classes (VOC; Bresilla et al., 2019). Pixel-wise accuracy was measured by comparison of ground truth and predicted information with the confusion matrix. Different levels of intersection over union (IoU) were considered, to evaluate the accuracy of our model (level of 0.4, 0.5, 0.6 in Tables 3, 4 and 5, respectively). The IoU computes the correctness of detection by calculating the overlapping area of prediction and the ground truth. Confusion matrices were generated, with the percentages of true positive, false negative and false positive. True negatives were not considered because this neural network wasn’t trained to distinguish or exclude different objects in the same picture.

		Predicted	
		True Positive (TP)	False Negative (FN)
ground truth (100%)	true	80.73%	19.26%
	false	False Positive (FP) 9.38%	True Negative (TN) 0%

133

Table 3. Confusion matrix with IoU = 0.4.

		Predicted	
		True Positive (TP)	False Negative (FN)
Ground truth (100%)	true	64.70%	35.29%
	false	False Positive (FP) 25.41%	True Negative (TN) 0%

134

135

Table 4. Confusion matrix with IoU = 0.5.

		Predicted	
		True Positive (TP)	False Negative (FN)
Ground truth (100%)	true	31.57%	68.42%
	false	False Positive (FP) 58.54%	True Negative (TN) 0%

136

Table 4. Confusion matrix with IoU = 0.6.

137 Considering the results, the threshold should be IoU = 0.5. Below this value the true positives can increase considerably  
138 but above it they decrease sharply, not reaching sufficient accuracy. The best F1 score of 0.84 was reached by imposing  
139 IoU=0.4. Recently, a multiple scale faster region-based convolutional neural network (MS-FRCNN) was used to detect  
140 small passion fruit by achieving 0.909 of the F1 score with IoU= 0.5 (Tu et al., 2020). Another framework was also  
141 presented by Stein et al., (2016) to identify, track, and localize mango using a faster region-based convolutional neural  
142 networks (R-CNN) detector with a F1 score of 0.89. Therefore, our results show that more images of pear boxes and  
143 more epochs to train our model are required, to estimate the number of pears. It is also important to highlight that fruit  
144 grown in the field generally have a different background when imaged that those sampled as the first layer of fruit in  
145 bins during storage. Consequently, the total number of fruit laid down can be estimated knowing the dimensions of the  
146 boxes.

### 147 3.2. Pear superficial scald models

148 Concerning the model for scald detection, two neural networks were trained (Fig. 2). The first one (Scald v1) was built  
149 starting from ten augmented datasets that represent five producers, four sampling points from 2019-2020 and one from  
150 2018-2019. It was then trained for one day up to one hundred epochs (approximately 100,000 iterations). Later, the

151 algorithm considering two confidence thresholds (CT) was tested and the confusion matrices were computed. In Table  
 152 5 and Table 6, V1 seems not to be very accurate but quite precise. In fact, the precision level was 0.62 for CT = 0.4 and  
 153 0.59 for CT = 0.2. On the other hand, the F1 scores do not reach enough reliability by achieving 0.21 with CT = 0.4  
 154 and 0.25 with CT = 0.2.

		Predicted	
		True	False
Ground Truth (100%)	True	True Positive (TP) 13.22%	False Negative (FN) 86.77%
	False	False Positive (FP) 8%	True Negative (TN) 0%

155 Table 5. Confusion matrix of scald V1 with IoU=0.5 and with confidence threshold =0.4

		Predicted	
		True	False
Ground Truth (100%)	True	True Positive (TP) 16.26%	False Negative (FN) 83.73%
	False	False Positive (FP) 11.1%	True Negative (TN) 0%

156 Table 6. Confusion matrix of scald V1 with IoU=0.5 and with confidence threshold =0.2

157 The second model (Scald v2) was trained starting from Scald v1 but adding 100 more epochs and part of the validation  
 158 dataset (the sixth producer). In any case, the images for the validation dataset (test set) there were not the same used for  
 159 the training dataset. Thereafter, the confusion matrices were extracted to evaluate the accuracy (Table 7 and Table 8).  
 160 Thus, the percentage of true positives improved for both confidence thresholds reaching almost 20 % with CT = 0.2.  
 161 Moreover, the F1 scores increase moderately from 0.21 to 0.26 in CT = 0.4 and from 0.25 to 0.28 in CT = 0.2.  
 162 Nevertheless, precision decreases and recall scores still low since the high number of false negatives in Scald v2.

		Predicted	
		True	False
Ground Truth (100%)	True	True Positive (TP) 16.71%	False Negative (FN) 83.28%
	False	False Positive (FP) 11.57%	True Negative (TN) 0%

163  
 164 Table 7. Confusion matrix of scald V2 with IoU=0.5 and with confidence threshold =0.4

		Predicted	
		True	False
Ground Truth (100%)	True	True Positive (TP) 19.17%	False Negative (FN) 80.82%
	False	False Positive (FP) 15.63%	True Negative (TN) 0%



Table 8. Confusion matrix of scald V2 with iou=0.5 and with confidence threshold =0.2

Looking at the scores and visual results of the models (Scald v1 and Scald v2), the algorithm can detect only few superficial scald symptoms. In fact, in many cases skin browning was not recognised. It was sometimes difficult to distinguish browning from a blemish or friction discoloration. However, evident symptoms of superficial scald were often detected (Fig. 2) and the YOLO v3 labels were quite precise (0.6 average precision). Comparing each pixel with its neighbour has proved accurate for the classification of apple fruit diseases and has achieved greater than 93% accuracy (Dubey & Jalal, 2012). In tomato plants some diseases, such as leaf mould, grey mould, canker, and plague, have shown variable performance and comparative results have shown that plain networks perform better than deeper networks. Fuentes et al., (2017) found that Faster R-CNN with VGG-16 had an average precision of 83%, compared to the same meta-architecture with ResNet-50, that achieved 75.37% or ResNeXt-50 with 71.1%. Thus, compared to the literature, the models can be improved carrying out four tasks: i) collect and ii) label more images to train the neural networks, iii) hire and teach qualified operators on how to label specific storage disorders and iv) finally apply different architectures such as the new versions of YOLO (v4 or v5) that deal with high resolution and more details or Mask R-CNN which detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance. A further action that could improve our results would be applying different image types in respect of a standard RGB images: converting these to other formats or colour-spaces such as grey scale or HSV applying pre-processing techniques.

### 3.3. Pear starch pattern index (SPI) models

Regarding the model for SPI recognition, different neural networks were trained as shown in Fig. 3. Tables 10 and 11 present the scores of different models and approaches to identify the right SPI. The CNN (YOLO v3) could not always detect the pear circle and the right maturity class. However, a remarkable increase was observed between the Ctifl and the Laimburg scales. The first reached 0.36 with the third version (Lugol v3) using a transfer learning method but the second (Lugol v5) achieved 0.59. Considering the F1 scores of each category, the class 1 in both scales and in all models is above 0.5. In the Laimburg scale the class 5 had the highest value. In pears, the recognition of the SPI is sometimes tricky for a trained operator as well. Nevertheless, in Rocha pear this analysis performed reasonably well and could be simply used in the field or at the laboratory (Saquet & Almeida, 2017). Decreasing the number of classes, the models based on the Laimburg scale were more precise and accurate, and, perhaps, more suitable for packing house needs. Considering the training method, transfer learning is effective to train SPI models. In this way, a prototype model could be initially trained with a few labelled pictures and less epochs. It could later be improved by adding images and epochs

195 until reasonable precision is obtained. The massive training session did not reach the same F1 score, especially with the  
 196 Ctfl scale. In fact, Lugol v4 had the same number of epochs as Lugol v3, but in this CNN the best checkpoint with the  
 197 lowest error (loss value) in each step was chosen.

	Transfer Learning			Massive Training
	Lugol V1 EP123 (EP150 In Total)	Lugol V2 EP130	Lugol V3 EP100	Lugol V4 EP305 (EP353 In Total)
1	0.6122	0.7599	0.6181	0.5333
2	0.3631	0.387	0.4126	0.2413
3	0.1558	0.2588	0.4042	0.2474
4	0.1818	0.2962	0.4	0.2735
5	0.198	0.1682	0.3174	0.3039
6	0.1801	0.1882	0.18	0.2
7	0.2368	0.1176	0.3578	0.2857
8	0.1904	0.1355	0.3478	0.1403
9	0.258	0.387	0.4615	0.5294
10	0.2	0.6153	0.3636	0.6153
Overall Average	0.2403	0.2791	0.3684	0.2939

198 Table 9. F1 scores of SPI models from 1 to 10 class (ctfl scale)

	Transfer Learning		Massive Training
	Lugol v5 EP124 (EP150 in total)	Lugol v6 EP 197 (EP229 in total)	Lugol v7 EP293 (EP353 in total)
1	0.72	0.6814	0.7121
2	0.5795	0.583	0.5763
3	0.5066	0.4705	0.4873
4	0.5813	0.5581	0.575
5	0.75	0.75	0.7843
Overall Average	0.5908	0.5717	0.5845

199 Table 10. F1 scores of SPI models from 1 to 5 class (Laimburg scale)

200 Confusion matrices were obtained with good results in both scales (Fig. 4 and Fig. 5). In fact, the highest number of  
 201 matches (true positives) were observed along the diagonal line. Thus, our CNNs can in general perform two goals at the  
 202 same time: detecting pears and identify the correct class. Concerning the recognition of the correct SPI, the models did  
 203 not make significant mistakes. In Figs. 4 and 5, in general one class of average error was found. Pear shape identification  
 204 was affected by a high number of false negatives. The neural networks still missed many objects: the model sometimes  
 205 did not recognize a fruit or even did not detect any fruit in a picture.

206 There is very little literature on objective tools to assess SPI and establish the ripening stage on apple or pear at harvest  
207 or during storage. However, Peirs et al., (2002) highlighted that the differences in SPI assessment carried out on the  
208 same samples from different subjects varied up to 60%. Furthermore, each estimator showed poor repeatability with  
209 respect to its own evaluation. Considering these elements, the use of an objective method would be necessary to improve  
210 accuracy, the reliability and repeatability of the evaluation of the SPI. For instance, Menesatti et al., (2009) investigated  
211 the use of hyper spectral NIR imaging to measure starch. This technique is not employed in the fruit supply chain  
212 because it does not properly consider the ripening stage and high costs as limiting factors. However, Zanella et al.,  
213 (2019) have used Amilon to quantify SPI as an effective tool. In ‘Golden Delicious’ apples, they showed a correlation  
214 and accuracy against the human evaluation of 0.94 and 0.58, respectively. However, currently the starch-meter needs a  
215 laboratory environment to work properly. By contrast, our algorithms could be implemented in portable device such as  
216 a standard smartphone, even if Lugol solution had to be previously applied on the pears.

#### 217 4. CONCLUSIONS

218 The possibility of identifying pears the first layer inside a box, to detect superficial scald on pear peel and to recognise  
219 SPI after the Lugol solution application, with high precision, using current deep neural network techniques and YOLO  
220 v3 as a convolutional neural network has been demonstrated. Results showed that further research is needed to improve  
221 the different models to recognise and quantify the level of superficial scald affecting fruit skin and to predict fruit  
222 storability considering SPI in pre- and post-harvest. More labelled pictures would improve the results; also, applying  
223 different approaches (e.g., image pre-processing/alteration) to increase the information extractable from the pictures by  
224 the ANN could boost considerably the algorithms. Our effort aims to help fruit cooperatives and producers improving  
225 the post-harvest management of pears to achieve consumer acceptance. They could adapt their storage techniques to  
226 identify and sort pear boxes in different cold rooms according to their ripening stage and symptoms of physiological  
227 disorders that are difficult to detect, even by trained human experts.

#### 228 ACKNOWLEDGMENT

229 We would like to thank all the students from University of Bologna that contributed to the project by labelling the  
230 images.

#### 231 REFERENCES

- 232 Agar, I. T., Biasi, W. V., & Mitcham, E. J. (1999). Exogenous ethylene accelerates ripening responses in Bartlett pears regardless of maturity or  
233 growing region. *Postharvest Biology and Technology*, 17(2), 67–78. [https://doi.org/10.1016/S0925-5214\(99\)00038-1](https://doi.org/10.1016/S0925-5214(99)00038-1)
- 234 Barreiro, P., Steinmetz, V., & Ruiz-Altisent, M. (1997). Neural bruise prediction models for fruit handling and machinery evaluation. *Computers*

235 *and Electronics in Agriculture*, 18(2–3), 91–103. [https://doi.org/10.1016/s1474-6670\(17\)47187-4](https://doi.org/10.1016/s1474-6670(17)47187-4)

236 Bazame, H. C., Molin, J. P., Althoff, D., & Martello, M. (2021). Detection, classification, and mapping of coffee fruits during harvest with  
237 computer vision. *Computers and Electronics in Agriculture*, 183(February). <https://doi.org/10.1016/j.compag.2021.106066>

238 Blanpied, G. D., Bramlage, W. J., Chu, C. L., Ingle, M., Kushad, M. M., Lau, O. L., & Lidster, P. D. (1991). A survey of the relationships  
239 among accumulated orchard hours below 10 °C, fruit maturity, and the incidences of storage scald on “Starkrimson Delicious” apples. *Canadian*  
240 *Journal of Plant Science*, 71(2), 605–608. <https://doi.org/10.4141/cjps91-091>

241 Bower, J. H., Biasi, W. V., & Mitcham, E. J. (2003). Effect of ethylene in the storage environment on quality of “Bartlett pears.” *Postharvest*  
242 *Biology and Technology*, 28(3), 371–379. [https://doi.org/10.1016/S0925-5214\(02\)00210-7](https://doi.org/10.1016/S0925-5214(02)00210-7)

243 Bresilla, K., Perulli, G. D., Boini, A., Morandi, B., Corelli Grappadelli, L., & Manfrini, L. (2019). Single-shot convolution neural networks for  
244 real-time fruit detection within the tree. *Frontiers in Plant Science*, 10(May), 1–12. <https://doi.org/10.3389/fpls.2019.00611>

245 Brouwer, B., Mensink, M., Hogeveen-van Echtelt, E., & Woltering, E. J. (2021). Pre-storage application of 1-methylcyclopropene does not  
246 affect the flavour of ‘Conference’ pears ripened after 8 months of commercial-standard controlled atmosphere storage. *Postharvest Biology and*  
247 *Technology*, 174(December 2020), 111448. <https://doi.org/10.1016/j.postharvbio.2020.111448>

248 Calvo, G., Candan, A. P., & Gomila, T. (2011). Post-harvest performance of “Abate Fetel” pears grown in Argentina in relation to harvest time.  
249 *Acta Horticulturae*, 909, 725–730. <https://doi.org/10.17660/ActaHortic.2011.909.89>

250 Calvo, Gabriela, Candan, A. P., Civello, M., Giné-Bordonaba, J., & Larrigaudière, C. (2015). An insight into the role of fruit maturity at harvest  
251 on superficial scald development in “Beurré D’Anjou” pear. *Scientia Horticulturae*, 192, 173–179. <https://doi.org/10.1016/j.scienta.2015.05.032>

252 Dubey, S. R., & Jalal, A. S. (2012). Detection and classification of apple fruit diseases using complete local binary patterns. *Proceedings of the*  
253 *2012 3rd International Conference on Computer and Communication Technology, ICCCT 2012*, 346–351.  
254 <https://doi.org/10.1109/ICCCT.2012.76>

255 ElMasry, G., Wang, N., & Vigneault, C. (2009). Detecting chilling injury in Red Delicious apple using hyperspectral imaging and neural  
256 networks. *Postharvest Biology and Technology*, 52(1), 1–8. <https://doi.org/10.1016/j.postharvbio.2008.11.008>

257 Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests  
258 recognition. *Sensors (Switzerland)*, 17(9). <https://doi.org/10.3390/s17092022>

259 Gama, R., Van Dyk, J. S., Burton, M. H., & Pletschke, B. I. (2017). Using an artificial neural network to predict the optimal conditions for  
260 enzymatic hydrolysis of apple pomace. *3 Biotech*, 7(2), 1–10. <https://doi.org/10.1007/s13205-017-0754-1>

261 Isidoro, N., & Almeida, D. P. F. (2006).  $\alpha$ -Farnesene, conjugated trienols, and superficial scald in “Rocha” pear as affected by 1-  
262 methylcyclopropene and diphenylamine. *Postharvest Biology and Technology*, 42(1), 49–56. <https://doi.org/10.1016/j.postharvbio.2006.05.003>

263 Jarvinen, T., Choi, D., Heinemann, P., & Baugher, T. (2018). Multiple Object Tracking-by-Detection for Fruit Counting on an Apple Tree  
264 Canopy. *ASABE Annual International Meeting*, 2–9. <https://doi.org/https://doi.org/10.13031/aim.201801193>

265 Jayas, D., & Karunakaran, C. (2005). Machine Vision System in postharvest technology. *Stewart Postharvest Review*, 1(2), 1–9. Retrieved from  
266 <http://www.portico.org/Portico/article?article=phx64r61jzw>

267 Kavdir, I., & Guyer, D. E. (2004). Comparison of artificial neural networks and statistical classifiers in apple sorting using textural features.  
268 *Biosystems Engineering*, 89(3), 331–344. <https://doi.org/10.1016/j.biosystemseng.2004.08.008>

269 Kim, J., Mowat, A., Poole, P., & Kasabov, N. (2000). Linear and non-linear pattern recognition models for classification of fruit from visible-  
270 near infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, 51(2), 201–216. [https://doi.org/10.1016/S0169-7439\(00\)00070-8](https://doi.org/10.1016/S0169-7439(00)00070-8)

271 Krizhevsky, A., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information*  
272 *Processing Systems*, 1097–1105.

273 Kupferman, E. (2001). Storage scald of apples. *Postharvest Information Network*, 1–7.

274 Leiva, G., Mondragon, G., Mery, D., & Aguilera, J. M. (2011). The automatic sorting using image processing improves postharvest blueberries  
275 storage quality. *Proceedings of International Congress on Engineering and Food (ICEF-2011)*. Retrieved from  
276 <http://web.ing.puc.cl/~dmery/Prints/Conferences/International/2011-ICEFa.pdf>

277 Lurie, S., & Watkins, C. B. (2012). Superficial scald, its etiology and control. *Postharvest Biology and Technology*, 65, 44–60.  
278 <https://doi.org/10.1016/j.postharvbio.2011.11.001>

279 Menesatti, P., Zanella, A., D'Andrea, S., Costa, C., Paglia, G., & Pallottino, F. (2009). Supervised multivariate analysis of hyper-spectral NIR  
280 images to evaluate the starch index of apples. *Food and Bioprocess Technology*, 2(3), 308–314. <https://doi.org/10.1007/s11947-008-0120-8>

281 Peirs, A., Scheerlinck, N., Perez, A. B., Jancsó, P., & Nicolai, B. M. (2002). Uncertainty analysis and modelling of the starch index during  
282 apple fruit maturation. *Postharvest Biology and Technology*, 26(2), 199–207. [https://doi.org/10.1016/S0925-5214\(02\)00038-8](https://doi.org/10.1016/S0925-5214(02)00038-8)

283 Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE*  
284 *Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>

285 Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement*. Retrieved from <http://arxiv.org/abs/1804.02767>

286 Saquet, A. A., & Almeida, D. P. F. (2017). Ripening physiology and biochemistry of ‘Rocha’ pear as affected by ethylene inhibition.  
287 *Postharvest Biology and Technology*, 125, 161–167. <https://doi.org/10.1016/j.postharvbio.2016.11.008>

288 Stein, M., Bargoti, S., & Underwood, J. (2016). Image based mango fruit detection, localisation and yield estimation using multiple view  
289 geometry. *Sensors (Switzerland)*, 16(11). <https://doi.org/10.3390/s16111915>

290 Sun, Y., Sun, H., Luo, M., Zhou, X., Zhou, Q., Wei, B., ... Ji, S. (2020). Membrane lipid metabolism in relation to core browning during  
291 ambient storage of ‘Nanguo’ pears. *Postharvest Biology and Technology*, 169(April), 111288.  
292 <https://doi.org/10.1016/j.postharvbio.2020.111288>

293 Tu, S., Pang, J., Liu, H., Zhuang, N., Chen, Y., Zheng, C., ... Xue, Y. (2020). Passion fruit detection and counting based on multiple scale faster  
294 R-CNN using RGB-D images. *Precision Agriculture*, 21(5), 1072–1091. <https://doi.org/10.1007/s11119-020-09709-3>

295 Vasconez, J. P., Delpiano, J., Vougioukas, S., & Auat Cheein, F. (2020). Comparison of convolutional neural networks in fruit detection and  
296 counting: A comprehensive evaluation. *Computers and Electronics in Agriculture*, 173(April), 105348.  
297 <https://doi.org/10.1016/j.compag.2020.105348>

298 Vasconez, Juan P., Kantor, G. A., & Auat Cheein, F. A. (2019). Human–robot interaction in agriculture: A survey and current challenges.  
299 *Biosystems Engineering*, 179, 35–48. <https://doi.org/10.1016/j.biosystemseng.2018.12.005>

300 Werth, K. (2000). *Colore & qualità delle mele dell'Alto Adige*. Consorzio delle cooperative ortofrutticole dell'Alto Adige (VOG).

301 Whitaker, B. D., Villalobos-Acuña, M., Mitcham, E. J., & Mattheis, J. P. (2009). Superficial scald susceptibility and  $\alpha$ -farnesene metabolism in  
302 “Bartlett” pears grown in California and Washington. *Postharvest Biology and Technology*, 53(1–2), 43–50.

303 <https://doi.org/10.1016/j.postharvbio.2009.04.002>

304 Zanella, A., Sadar, N., Ebner, I., & Panarese, A. (2019). Amidometro, strumento per valutare lo stato di maturazione delle mele. *Rivista Di*  
305 *Frutticoltura e Di Ortofrutticoltura*. Retrieved from <https://rivistafrutticoltura.edagricole.it/post-raccolta/amidometro-strumento-per-valutare-lo->  
306 [stato-di-maturazione-delle-mele/](https://rivistafrutticoltura.edagricole.it/post-raccolta/amidometro-strumento-per-valutare-lo-stato-di-maturazione-delle-mele/)

## 307 **FIGURE CAPTIONS**

308 Fig. 1. Training graph of Lugol V3 model where the purple bars are the training rate, and the orange line is the loss value in each epoch (extracted  
309 by <https://supervise.ly>).

310 Fig. 2. Example of superficial scald detection (polygon is ground truth, rectangle is Scald v2 detection).

311 Fig. 3. Example of starch index recognition (one rectangle is real ground truth, the other one is predicted by LUGOL v3. Matching colours means  
312 correct classification while different colours is wrong classification. Missing double rectangle is missed detection.

313 Fig. 4. Confusion Matrix of Lugol v3 with IoU = 0.5 and confidence threshold = 0.5.

314 Fig. 5. Confusion Matrix of Lugol v5 with IoU = 0.5 and confidence threshold = 0.5.