

Research Paper

A quality decay model with multinomial logistic regression and image-based deep learning to predict the firmness of ‘Conference’ pears in the downstream supply chains

Xuezhen Guo^{*}, Aneesh Chauhan, Jan Verschoor, Andrei Margert

Wageningen Food & Biobased Research, the Netherlands

ABSTRACT

Traditional quality-decay models (e.g., multinomial logistic regression) for fruit quality classification deals with tabular data which focus mainly on the storage parameters such as storage duration and conditions (D&C). Those parameters have the effects on quality decay at an aggregate scale across different experimental levels; they are not good at capturing the variations within each experimental level. This may restrict the predictive power of the traditional model. On the contrary, image-based deep learning models are dealing with individual products and can extract the deep features of each fruit to provide individual-based quality information but lack information regarding the post-harvest conditions (time of harvest, storage conditions etc.).

In this research, we investigate the combined performance of the multinomial logistic regression with the image-based convolutional neural network (CNN) for quality prediction of ‘Conference’ pears (*Pyrus communis* L.) (measured by firmness) where the extracted deep features are used as the explanatory variables for the logistic regression model. The results show that combining deep features with D&C parameters are likely to improve the predictive power of the logistic regression model to predict the firmness of the conference pears. The managerial implications as well as future research directions are also discussed.

1. Introduction

Fruits and vegetables are perishable products which account roughly for 50% of the global food loss and waste (FLW)(Guo et al., 2020). Fruits and vegetables are also the major nutrient sources for some critical nutrients (e.g., vitamins). Therefore, reducing fruit and vegetable waste have significant implications to food and nutrition security. To reduce fruits and vegetable wastes, an important step is to understand the quality-decay process of the products, which enable us to predict the product qualities in the postharvest supply chains. There are different types of quality decay models for fruits and vegetables available in the literature. Mechanistic models are the most popular ones which specify the decaying processes by a couple of ordinary differential equations (e.g., Defraeye et al., 2019; Jalali et al., 2021; Jalali et al., 2020; Nourian et al., 2003; Piagentini et al., 2005; Schouten et al., 1997; Schouten et al., 2002; Schouten et al., 2009; Tijskens and Polderdijk, 1996; Villanueva et al., 2005). Mechanistic models are considered to have the difficulty in capturing the heterogeneity of the fruits within a batch. As a result, generalized linear models such as logistic regressions have been used as an alternative for quality indicator prediction (e.g., Lammertyn et al., 2000; Marini et al., 2022; Tolesa et al., 2018). Another stream of widely used models to predict fruit quality are the ones exploiting the

hyperspectral imaging data with machine or deep learning technologies (e.g., Li et al., 2017; Wang et al., 2016; Xiao and Li, 2022; Zhu et al., 2017). Since such models extract the spectral profile of individual products, they are considered as good at capturing product-specific variations.

Firmness is an important indicator to reflect the quality of the fruit products (Diezma-Iglesias et al., 2006; Fathizadeh et al., 2021; Tor-kashvand et al., 2017), especially from the storability point of view. There are specific studies focusing on developing prediction models to predict fruit firmness. Hernández et al. (2021) developed a mechanistic model to predict the softening behavior of the Chilean ‘Hass’ avocados and claim it has the potential to forecast the change of avocado’s firmness when the distribution of the initial single enzyme complex responsible for the breakdown of firmness is known. Hertog et al. (2003) investigated the softening process of ‘Hass’ avocados in different modified atmosphere (MA) conditions with a mechanistic model that links the firmness to the rate of gas changes. Gwanpua et al. (2013) developed a kinetic model that specifies the firmness decay of three apple cultivars as a function of time, temperature, CA conditions and endogenous ethylene concentration. Hertog et al. (2016) further modified the model to predict the firmness of kiwifruit with exogenous ethylene levels.

^{*} Corresponding author.

E-mail address: xuezhen.guo@wur.nl (X. Guo).

<https://doi.org/10.1016/j.jspr.2024.102450>

Received 12 May 2024; Received in revised form 13 October 2024; Accepted 18 October 2024

Available online 25 October 2024

0022-474X/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1
Storage duration and condition

Storage D&C	Description
harvest	Firmness is measured at moment of harvest
harvest + 1 wk 10 °C	Firmness is measured one week after the harvest with the storage temperature 10 °C
harvest + 1 wk 18 °C	Firmness is measured one week after the harvest with the storage temperature 18 °C
after 6 months CA ^a	Firmness is measured after 6 months of CA storage
after 9 months CA	Firmness is measured after 9 months of CA storage
after 9 months CA + 1 wk 10 °C	Firmness is measured after 9 months of CA storage plus 1 week storage under 10 °C

^a The CA condition is "temperature −0.5 °C, with 3% O₂ and ≤0.8 % CO₂".

Next to the models that utilize the storage conditions data to predict fruit firmness, another stream of literature tries to employ the hyperspectral imaging data to identify the softening status of the fruits. Examples include Lu and Peng (2006) for peaches, Mendoza et al. (2011) and Peng and Lu (2008) for apples, Mishra et al. (2021) for avocados. Xiao and Li (2022) is the first to combine hyperspectral data with the postharvest storage condition data to predict the firmness of the kiwi-fruits and find it can improve accuracy of the prediction. Other recent studies on using hyperspectral imaging data to predict fruit quality include (e.g., Mishra et al., 2022; Qi et al., 2023; Sun et al., 2024)

Although hyperspectral data can provide precise phenotyping information on the fresh product for firmness prediction, it requires specific equipment to perform the hyperspectral scanning of the individual products. Such scanning may be relatively easily to apply at some stages of the supply chains (e.g., at the sorting line) but do have the difficulty for application in the downstream of the supply chains (e.g., at the retail stage). Therefore, a more convenient method with the easy-to-get data needs to be developed to predict fruit firmness in the downstream fruit

supply chains.

In this research, we propose a novel model that combines the multi-nominal logistic regression and deep learning with camera image data to predict the firmness of the ‘Conference’ pears. The deep learning model extracted the deep features of the ‘Conference’ pears and use them as the explanatory variables for the multi-nominal logistic regression. The benefit of applying logistic regression is its flexibility of incorporating more variables and its ability of capturing the variations of the products (Lammertyn et al., 2000).

2. Materials and method

In this research, we use the data of ‘Conference’ pears through combining the datasets from different experiments and different years, provided by multiple Dutch ‘Conference’ pear commercial orchards in the seasons 2015–2017. The firmness of the pears were measured under 6 different storage duration & condition (D&C) (listed in Table 1)

Since the purpose of this study is to demonstrate the validity and feasibility of improved firmness prediction by combining camera images with storage condition data, we will treat every data point equally important for parameter estimation without taking into account, e.g., the orchard effects, year effects and other unobserved effects. This is in line with the approach applied by Lammertyn et al. (2000).

The firmness of the ‘Conference’ pears is measured by force (in kg) with the penetrometer. Since decision making in the fresh produce supply chains usually require classification rather than regression (Xiao and Li, 2022), we converted firmness from a continuous variable to a categorical variable. For the explanatory variables to predict the firmness, there are six storage durations & conditions (D&C) applied to the ‘Conference’ pears (see Table 1), each of which is modelled as a binary independent variable of the multi-nominal logistic regression model.

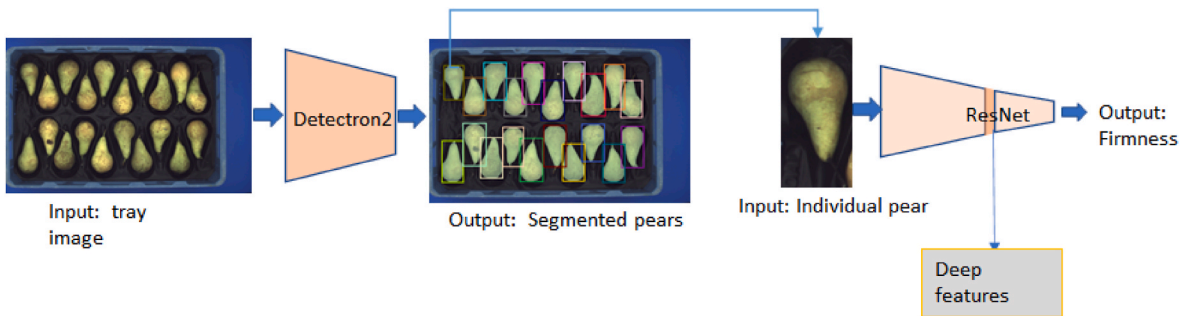


Fig. 1. Conceptual view of the image processing and deep feature extraction pipeline.

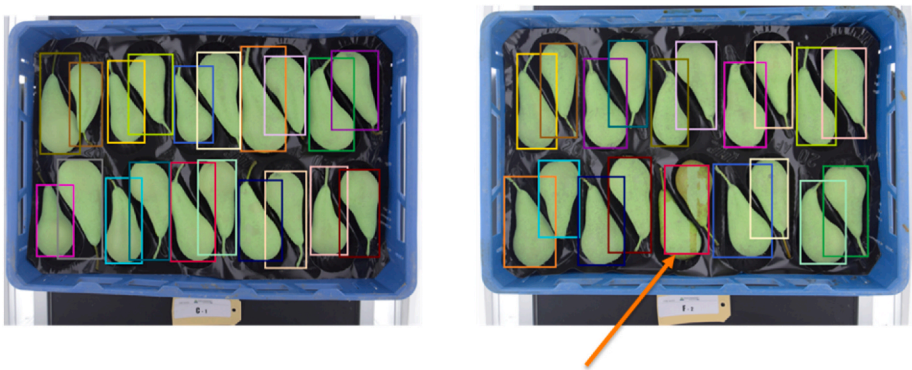


Fig. 2. The bounding boxes and masks for each detected pear. Image on the left demonstrates an example where the detection was unsuccessful, and on the right an example where all pears are detected successfully.

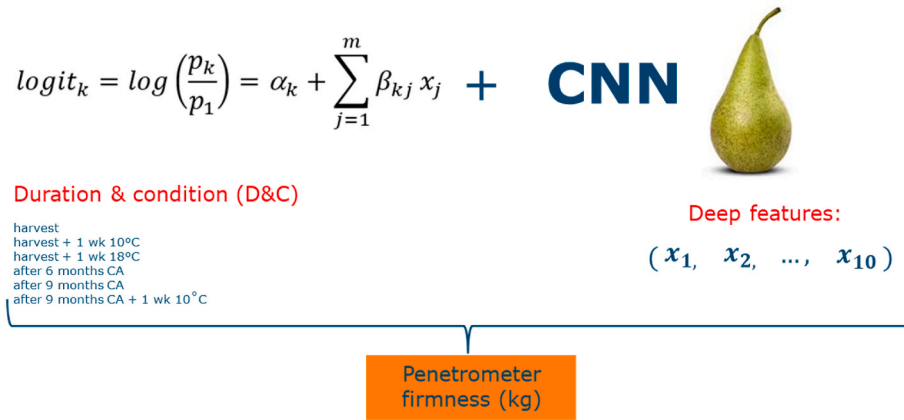


Fig. 3. Schematic representation of the modelling approach.

Table 2

The prediction accuracy of the model on the testing data for different computational experiments with 3 firmness classes.

Experiment	Prediction accuracy with 3 firmness classes	Prediction accuracy with 6 firmness classes
D&C	0.73	0.48
Deep features	0.7	0.5
D&C + Deep features	0.77	0.54

For extracting visual features, Fig. 1 demonstrates the conceptual data processing pipeline. Since the fruits were stored in the crates, a part of the visual processing involved individual fruit extraction. For that purpose, Mask-RCNN (He et al., 2017) deep neural network from the Detectron 2 software library (Wu et al., 2019) was trained. Fig. 2 shows a couple trays demonstrating the detections. For a very few crates there were mistakes in detection, and those images are discarded in the next step of data processing.

From the extracted individual pear images, ResNet (He et al., 2017) network was modified and trained to predict pear firmness. A key modification to the network is that an additional fully-connected layer of 10 nodes was added to the network right before the output layer. It from this layer the deep-features are extracted. The details for the ResNet model are presented in the Appendix.

The logit function is also called log-odds function which transfers the binary response variable to its log odds form. The log-odds function can then be predicted by a linear combination of the independent variables. In the multinomial logistic regression context, the logit function is specified in equation (1):

$$\text{logit}_k = \log\left(\frac{p_k}{p_1}\right) = \alpha_k + \sum_{j=1}^m \beta_{kj} x_j \quad (1)$$

where:

k indicates the k th level/class of response variable (where $k = 1$ is the reference class)

p_k is the probability of a ‘Conference’ pear belonging to class k

α_k is the intercept for the k th class

x_j is the explanatory variable j

β_{kj} is the regression coefficient for explanatory variable j and class k

One point needs to be pointed out here is that the logistic regression model in this research is slightly different from the model in Lammertyn et al. (2000). Lammertyn et al. (2000)’s model has grouped data with aggregated percentages for different classes in each batch. As a result, the individual-specific values are eliminated, which means loss of information on individual heterogeneities and therefore less preferred compared to the model developed in this research. Moreover, since Lammertyn et al. (2000)’model also does not take into account the unobserved batch-specific effects, it essentially treat all the individual pears equal in parameter estimation. This makes their model still largely comparable to the model developed in this paper. The schematic representation of our modelling approach is presented in Fig. 3.

In the context of fruit quality, converting the continuous variable “firmness” into a categorical variable is in line with the industrial practice and can help to better understand and predict the quality status of the fruits. By dividing the continuous variable into several classes, we can identify the different levels of firmness in the fruits and assess whether they meet certain quality standards. In the case of the ‘Conference’ pears, we have chosen to use three distinct classes for the logistic regression model, namely “ ≤ 3 ”, “ $3 < f \leq 5$ ” and “ $f > 5$ ”. This decision was based on the current industry standards and practices, where these three classes are commonly used to categorize the firmness of the fruits.

However, it is worth noting that the number of classes used in the categorical variable can impact the accuracy and reliability of the logistic regression model. In some cases, using more classes can provide more detailed information on the quality status of the fruits, but it can also increase the complexity of the model and reduce its predication accuracy. Therefore, in addition to the three-class model, we have also developed a six-class model that includes the classes “ ≤ 2 ”, “ $2 < f \leq 3$ ”, “ $3 < f \leq 4$ ”, “ $4 < f \leq 5$ ”, “ $5 < f \leq 6$ ”, and “ $f > 6$ ”. This model offers a more nuanced categorization of the firmness levels of the fruits, which can be particularly useful in applications where a finer-grained assessment of

Table 3

The confusion matrixes for the experiments with 3 classes.

D&C				Deep features				D&C + Deep features			
	Reference				Reference				Reference		
Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$	Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$	Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$
$f \leq 3$	142	19	0	$f \leq 3$	120	33	10	$f \leq 3$	155	14	2
$3 < f \leq 5$	0	0	0	$3 < f \leq 5$	27	49	47	$3 < f \leq 5$	8	54	60
$f > 5$	21	137	343	$f > 5$	16	64	296	$f > 5$	1	66	302

Table 4
The confusion matrixes for the experiments with 6 classes.

D&C													
Deep features							D&C + Deep features						
Reference							Reference						
Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$	Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$
$f \leq 2$	0	0	0	0	0	0	$f \leq 2$	32	17	4	7	6	3
$2 < f \leq 3$	70	75	11	0	0	0	$2 < f \leq 3$	31	52	18	16	7	2
$3 < f \leq 4$	0	0	0	0	0	0	$3 < f \leq 4$	1	3	1	2	0	0
$4 < f \leq 5$	0	0	0	0	0	0	$4 < f \leq 5$	12	14	4	31	30	1
$5 < f \leq 6$	13	18	11	126	245	93	$5 < f \leq 6$	4	8	4	54	205	72
$f > 6$	0	0	0	0	0	0	$f > 6$	0	0	0	2	10	9
							Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$
							$f \leq 2$	45	23	2	1	0	1
							$2 < f \leq 3$	39	68	13	2	0	0
							$3 < f \leq 4$	4	2	4	3	0	0
							$4 < f \leq 5$	1	0	5	35	39	6
							$5 < f \leq 6$	0	1	8	70	200	69
							$f > 6$	0	1	0	3	11	6

Table 5
The prediction accuracy of the model on the testing data for different computational experiments with 3 firmness classes with color index added.

Experiment	Prediction accuracy with 3 firmness classes	Prediction accuracy with 6 firmness classes
D&C	0.75	0.5
D&C + Deep features	0.78	0.54
D&C + Color index	0.79	0.5
D&C + Deep features + Color index	0.78	0.56

fruit quality is required. Overall, the choice of the number of classes in the categorical variable should depend on the specific application and the level of detail required in the quality assessment. By incorporating the categorical variable into the logistic regression model, we can better predict the firmness of the ‘Conference’ pears and provide valuable information to optimize the post-harvest supply chains.

The overall dataset (in the format of a dataframe) was constructed by integrating multiple experimental data, which contains 7812 pears.

In step 1, for the deep learning model, we first split the overall dataset into the training and inference datasets using an 80%–20% ratio, which results in 6250 and 1562 pears in the training and inference datasets respectively. After the training of the deep learning model, we use the trained model to do inference for the 1562 pears: all the learned weights are kept the same during the inference process. In this way we obtained the deep features for the pears in both the training and inference datasets and append them the original dataframe.

In step 2, we randomly select 900 pears from both the training and the inference datasets and combine them into a training dataset (with 1800 pears) to train the logistic regression model. The remaining 662 pears in the inference dataset are served as the testing dataset for the logistic regression model. The reason for selecting the pears from both training and inference datasets to train logistic regression is as follows. The deep learning model tries to fit the training data as much as possible (even though with the L1 regularization). If we only use the training dataset in step 1 to train the logistic regression model (and using the reference dataset in step 1 as the testing dataset), the learned features, purely based on the training dataset of the deep learning model, will be somewhat misleading in predicting the response variable in the testing dataset. Incorporating a random subset of datapoints from the reference dataset to train the logistic regression model essentially calibrates those deep features with logistic regression’s coefficients. This actually complies with the meaning of hybrid modelling, which combines the strengths of different modelling techniques.

3. Results

3.1. The results of experiments with storage durations & conditions (D&C) and deep features

We start with the results of experiments with the analyses on D&C and deep features (shown in Table 2). For the model works with 3 firmness classes, the prediction accuracy using D&C as the only predictor, the prediction accuracy is 0.73. If only the deep features are used as predictors, the prediction accuracy decreases to 0.7. Finally, if we combine D&C and deep features, the prediction accuracy grows to 0.77.

When the model deals with 6 firmness classes, first of all, the prediction accuracies in general decrease compared to the situation under the 3 firmness classes. This makes sense because as the granularity of the response variable increases, the prediction accuracy is logically reduced. Even though the absolute value decreases, the pattern of the accuracies under different experiments is still noteworthy. The combination of D&C and deep features as predictors (0.54) outperforms the models using each of them individually (0.48 and 0.5).

Table 3 shows different confusion matrixes for the experiments with 3 firmness classes. It looks D&C has the difficulty in correctly

Table 6

The confusion matrixes for the experiments with 3 classes with color index added.

D&C				D&C + Deep features				D&C + Color index				D&C + Deep features + Color index			
Reference				Reference				Reference				Reference			
Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$	Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$	Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$	Prediction	$f \leq 3$	$3 < f \leq 5$	$f > 5$
$f \leq 3$	107	9	0	$f \leq 3$	103	9	0	$f \leq 3$	107	9	0	$f \leq 3$	103	9	0
$3 < f \leq 5$	1	32	40	$3 < f \leq 5$	5	25	22	$3 < f \leq 5$	1	29	27	$3 < f \leq 5$	5	24	21
$f > 5$	0	15	57	$f > 5$	0	22	75	$f > 5$	0	18	70	$f > 5$	0	23	76

Table 7

The confusion matrixes for the experiments with 6 classes with color index added.

D&C							D&C + Deep features						
Reference							Reference						
Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$	Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$
$f \leq 2$	61	52	6	0	0	0	$f \leq 2$	43	21	1	0	0	0
$2 < f \leq 3$	0	0	0	0	0	0	$2 < f \leq 3$	16	29	4	0	0	0
$3 < f \leq 4$	0	0	0	0	0	0	$3 < f \leq 4$	2	3	2	3	0	0
$4 < f \leq 5$	0	1	7	27	30	9	$4 < f \leq 5$	0	0	5	12	16	4
$5 < f \leq 6$	0	0	0	3	37	15	$5 < f \leq 6$	0	0	0	14	48	19
$f > 6$	0	0	0	0	0	0	$f > 6$	0	0	1	1	3	1

D&C + Color index							D&C + Deep features + Color index						
Reference							Reference						
Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$	Prediction	$f \leq 2$	$2 < f \leq 3$	$3 < f \leq 4$	$4 < f \leq 5$	$5 < f \leq 6$	$f > 6$
$f \leq 2$	61	52	6	0	0	0	$f \leq 2$	44	21	1	0	0	0
$2 < f \leq 3$	0	0	0	0	0	0	$2 < f \leq 3$	15	29	4	0	0	0
$3 < f \leq 4$	0	0	0	0	0	0	$3 < f \leq 4$	2	3	3	3	0	0
$4 < f \leq 5$	0	1	7	20	23	3	$4 < f \leq 5$	0	0	4	13	15	3
$5 < f \leq 6$	0	0	0	10	44	21	$5 < f \leq 6$	0	0	0	13	49	20
$f > 6$	0	0	0	0	0	0	$f > 6$	0	0	1	1	3	1

distinguishing the medium firmness class of pears ($3 < f \leq 5$) from the ones with the best and worst firmness classes ($f \leq 3$ and $f > 5$). Adding deep features as additional predictors seems to fix the problem to some extent because deep features are by nature good at capturing the granularity in the data. Furthermore, the model that integrates both D&C and deep features demonstrates better performance compared to those utilizing either one as predictors independently; this may suggest that the combined model leverage the strengths of diverse data types.

A similar pattern can be found for the 6-class experiments in Table 4 as well. If only using D&C as predictors, the model classified all pears to 2 classes ($2 < f \leq 3$ and $5 < f \leq 6$) while adding deep feature can somewhat mitigate this issue.

3.2. The results of experiments with D&C, deep features and color index

Part of the dataset (used in Section 3.1) also included the measurements of color index for individual pears with a scale from 1 to 5. In this section, we use this color-indexed sub-dataset to retrain and retest the model where color index is added as an additional predictor. We do this to especially compare its added value with that of the deep features.

Table 5 shows the prediction accuracy of the model on the new testing dataset for four different model settings. Similar to the results in section 3.1, for the 3-classes experiments, adding deep features and/or color index to D&C improves the prediction accuracy of the model. Adding color index performs slightly better than adding deep features but it can well be the effect of data randomness. For the 6-class experiments, adding deep features improves the prediction accuracy and further adding color index increase the accuracy more.

If we look into the confusion matrices, for the 3-class experiments, the adding of deep features and color index seems to especially improve the performance of correctly classifying class $f > 5$ at the cost of misclassifying somewhat more class $3 < f \leq 5$ (Table 6).

When referring to the experiments under 6 firmness classes, we can

see that adding color index does not help to improve the classification for certain classes while adding deep features does (Table 7). This implies that deep features can capture the granular differences between more trivial firmness classes while color index can not.

4. Discussion

In this paper, we have investigated the feasibility of combining deep features of ‘Conference’ pears with the storage D&C to predict their firmness, employing multinomial logistic regression. The deep features are extracted by the CNN model RestNet on the camera image data. The extracted deep features are then used as predictors for the logistic regression model. In addition to incorporating deep features, we further extend our analysis by including color index as a predictor to assess its comparative efficacy to deep features. The response variable, firmness, is segmented for 3 and 6 classes to test the model’s performances under different measurement granularities.

The results show that using the deep features as additional information sources can indeed improve the prediction accuracy of the model compared to the model only using the D&C as predictors. This makes sense because D&C are applied across all classes of pears, which by nature has the difficulty in differentiating the granular differences among individuals. Adding color index seems to have similar effects as the deep features on the differentiation of pears for 3-firmness classes. However, it has the difficulty in improving the classification accuracy for the 6-firmness class scenario. It implies that color index can capture somewhat granular differences between individual pears, however with a less extent compared to deep features extracted by the deep learning model.

Although cross-study comparisons are challenging due to the lack of similar research utilizing camera images under comparable conditions, we have compared our results with two studies that employed vis-NIR spectroscopy. The prediction accuracies we obtained for the 3-class

problem range from 0.7 to 0.8, which is higher than the 0.66 reported by Wang et al. (2017) but lower than the 0.93 reported by Mishra et al. (2022). The lower accuracy in our case, compared to Mishra et al. (2022), is expected, as vis-NIR spectroscopy is inherently more precise than camera imaging or color indices for predicting firmness.

Several managerial implications can be derived from this research. Firstly, in practice, companies in the fruit chains (e.g., packing houses, wholesalers, retailers) may consider to combine the information from fruit-storage history and camera-image based deep learning to improve the prediction accuracy on fruit firmness. Since the costs associated with gathering image using camera are relatively low, it is a promising cost-effective method to predict fruit firmness information. Secondly, deep features can outperform color index with respect to the improved firmness prediction accuracy when the granularity of the response variable is high. This means the relatively costly color index measurement can be replaced by the deep features learned from fruit images, especially when more types of fruits need to be classified. Those can significantly improve the profitability of the relevant companies in the fruit supply chains.

We should also not ignore the limitation of this study especially considering the limited variations in the current dataset. Future research should focusing on gathering the data with more biological and managerial variations (e.g., different softening rates and storage conditions). Furthermore, given the rapid advancements in AI, it is essential to stay updated on emerging AI models that could further enhance prediction accuracy. Finally, the model should be evaluated in a real-world setting

to determine its practical value and potential impact on actual business operations.

CRedit authorship contribution statement

Xuezheng Guo: Writing – original draft, Validation, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Aneesh Chauhan:** Writing – review & editing, Validation, Supervision, Investigation, Formal analysis, Conceptualization. **Jan Verschoor:** Writing – review & editing, Data curation. **Andrei Margert:** Writing – review & editing, Formal analysis.

Declaration of competing interest

We hereby declare that there are no conflicts of interests for our research on this paper.

Acknowledgment

The authors would like to acknowledge the funding from the Wageningen University & Research "Data Driven & High Tech (DDHT) program (KB-38) with the project number KB-38-001-003" that is supported by the Ministry of Agriculture, Fisheries, Food Security and Nature.

Appendix

Deep Features are represented by a set of unbound numeric values. Such set is calculated for each pear using its colour image and ResNet convolutional neural network (CNN). In a nutshell, ResNet is composed of a series of convolutional blocks followed by a single fully connected (dense) layer. Below is

the equation expressing the final layer in the ResNet forward pass after all the convolution layers and anverage pooling:

$$\hat{y} = \sigma(x_{ap} W_{fc}) \quad (1)$$

where $\hat{y}$ is the predicted target value, σ — softmax activation function, x_{ap} — average pooling operation output in a forward pass, W_{fc} — fully connected layer weights. Deep Features are calculated by passing a pear image forward through the modified ResNet 34 (number 34 indicates the depth of ResNet, 34 layers). The table below, from the original ResNet paper (He et al., 2017), indicates architecture differences between ResNet variants.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Convolutional blocks remain unaltered, whereas the final fully connected layer is replaced by two new fully connected layers. The equation below, similarly to equation (1), expresses the operations after all the convolution blocks and average pooling:

$$\hat{y} = \sigma(x_{ap} W_1 W_2) \quad (2)$$

where W_1 is the 512-by-10 wt matrix, and W_2 — 10-by-1 weight matrix. The result of $x_{ap}W_1$ is what we are calling the Deep Features: a set of 10 values extracted before the last fully connected network layer of the modified ResNet 34.

After this change, our version of the network output is just a single value as target prediction. Target is set to be pear firmness obtained in the experiment with penetrometer. Pear firmness is characterized by a real number ranging from 0 to infinity.

This modified ResNet is trained on the dataset of images, in total containing 7812 different pears. Each pear is treated as a single data point, meaning the modified ResNet yields firmness predictions for

each individual fruit. Data split is random: 6250 pear for training and 1562 for validation. Random algorithms are initialized with the fixed seed. Data points' target values are not considered when splitting the dataset.

Images are resized to 256-by-256 pixels. Moreover, pixel values are normalized with the following parameters per colour (red, green, blue):

- color mean = [0.485, 0.456, 0.406].
- color standard deviation = [0.229, 0.224, 0.225],

Both training loss and validation metric are L1 loss, Both, training loss and validation metric are L1 loss (Mean Absolute Error). Batch [Not Available in CrossRef] [Not Available in Internal Pubmed] [Not Available in External Pubmed]

size, size, learn rate, and weight decay are subjects of hyperparameter optimization. [Not Available in CrossRef] [Not Available in Internal Pubmed] [Not Available in External Pubmed]

See the numbers below to gain a high level overview of the explored hyperparameter.

Combinations:

- starting learn rate is 0.0003019943053251939,
- learn rate scheduler is the default pytorch ReduceOnPlateau,
- optimizer is Adam,
- batch size is 69,
- weight decay is 1.1085777591277749e-07.

Data availability

The authors do not have permission to share data.

References

- Defraeye, T., Tagliavini, G., Wu, W., Prawiranto, K., Schudel, S., Assefa Kerisima, M., Verboven, P., Bühlmann, A., 2019. Digital twins probe into food cooling and biochemical quality changes for reducing losses in refrigerated supply chains. *Resour. Conserv. Recycl.* 149, 778–794. <https://doi.org/10.1016/j.resconrec.2019.06.002>.
- Diezma-Iglesias, B., Valero, C., García-Ramos, F.J., Ruiz-Altisent, M., 2006. Monitoring of firmness evolution of peaches during storage by combining acoustic and impact methods. *J. Food Eng.* 77 (4), 926–935.
- Fathizadeh, Z., Aboonajmi, M., Hassan-Beygi, S.R., 2021. Nondestructive methods for determining the firmness of apple fruit flesh. *Information Processing in Agriculture* 8 (4), 515–527. <https://doi.org/10.1016/j.inpa.2020.12.002>.
- Guo, X., Broeze, J., Groot, J.J., Axmann, H., Vollebregt, M., 2020. A worldwide hotspot analysis on food loss and waste, associated greenhouse gas emissions, and protein losses. *Sustainability* 12 (18), 7488.
- Gwanpua, S.G., Verlinden, B.E., Hertog, M.L.A.T.M., Van Impe, J., Nicolai, B.M., Geeraerd, A.H., 2013. Towards flexible management of postharvest variation in fruit firmness of three apple cultivars. *Postharvest Biol. Technol.* 85, 18–29. <https://doi.org/10.1016/j.postharvbio.2013.04.015>.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969.
- Hernández, I., Uarota, V., Paredes, D., Fuentealba, C., Defilippi, B.G., Campos-Vargas, R., Meneses, C., Hertog, M., Pedreschi, R., 2021. Can metabolites at harvest be used as physiological markers for modelling the softening behaviour of Chilean “Hass” avocados destined to local and distant markets? *Postharvest Biol. Technol.* 174, 111457.
- Hertog, M.L.A.T.M., Jeffery, P.B., Gwanpua, S.G., Lallu, N., East, A., 2016. A mechanistic model to describe the effects of time, temperature and exogenous ethylene levels on softening of kiwifruit. *Postharvest Biol. Technol.* 121, 143–150. <https://doi.org/10.1016/j.postharvbio.2016.08.002>.
- Hertog, M.L.A.T.M., Nicholson, S.E., Whitmore, K., 2003. The effect of modified atmospheres on the rate of quality change in ‘Hass’ avocado. *Postharvest Biol. Technol.* 29 (1), 41–53. [https://doi.org/10.1016/S0925-5214\(02\)00211-9](https://doi.org/10.1016/S0925-5214(02)00211-9).
- Jalali, A., Linke, M., Geyer, M., Mahajan, P., 2021. Integrative programming for simulation of packaging headspace and shelf life of fresh produce. *MethodsX* 8, 101514.
- Jalali, A., Linke, M., Geyer, M., Mahajan, P.V., 2020. Shelf life prediction model for strawberry based on respiration and transpiration processes. *Food Packag. Shelf Life* 25, 100525. <https://doi.org/10.1016/j.fpsl.2020.100525>.
- Lammertyn, J., Aerts, M., Verlinden, B.E., Schotsmans, W., Nicolai, B.M., 2000. Logistic regression analysis of factors influencing core breakdown in ‘Conference’ pears. *Postharvest Biol. Technol.* 20 (1), 25–37.
- Li, M., Pullanagari, R.R., Pranamornkith, T., Yule, I.J., East, A.R., 2017. Quantitative prediction of post storage ‘Hayward’ kiwifruit attributes using at harvest Vis-NIR spectroscopy. *J. Food Eng.* 202, 46–55.
- Lu, R., Peng, Y., 2006. Hyperspectral scattering for assessing peach fruit firmness. *Biosyst. Eng.* 93 (2), 161–171.
- Marini, R.P., Lavelly, E.K., Baugher, T.A., Crassweller, R., Schupp, J.R., 2022. Using logistic regression to predict the probability that individual ‘Honeycrisp’ apples will develop bitter pit. *Hortscience* 57 (3), 391–399.
- Mendoza, F., Lu, R., Ariana, D., Cen, H., Bailey, B., 2011. Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biol. Technol.* 62 (2), 149–160.
- Mishra, P., Brouwer, B., Meesters, L., 2022. Improved understanding and prediction of pear fruit firmness with variation partitioning and sequential multi-block modelling. *Chemometr. Intell. Lab. Syst.* 222, 104517.
- Mishra, P., Paillart, M., Meesters, L., Woltering, E., Chauhan, A., Polder, G., 2021. Assessing avocado firmness at different dehydration levels in a multi-sensor framework. *Infrared Phys. Technol.* 118, 103901.
- Nourian, F., Ramaswamy, H.S., Kushalappa, A.C., 2003. Kinetics of quality change associated with potatoes stored at different temperatures. *LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.)* 36 (1), 49–65. [https://doi.org/10.1016/S0023-6438\(02\)00174-3](https://doi.org/10.1016/S0023-6438(02)00174-3).
- Peng, Y., Lu, R., 2008. Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. *Postharvest Biol. Technol.* 48 (1), 52–62.
- Piagentini, A.M., Mendez, J.C., Guemes, D.R., Pirovani, M.E., 2005. Modeling changes of sensory attributes for individual and mixed fresh-cut leafy vegetables. *Postharvest Biol. Technol.* 38 (3), 202–212.
- Qi, H., Shen, C., Chen, G., Zhang, J., Chen, F., Li, H., Zhang, C., 2023. Rapid and non-destructive determination of soluble solid content of crown pear by visible/near-infrared spectroscopy with deep learning regression. *J. Food Compos. Anal.* 123, 105585.
- Schouten, R.E., Otma, E.C., van Kooten, O., Tijskens, L.M.M., 1997. Keeping quality of cucumber fruits predicted by biological age. *Postharvest Biol. Technol.* 12 (2), 175–181. [https://doi.org/10.1016/S0925-5214\(97\)00045-8](https://doi.org/10.1016/S0925-5214(97)00045-8).
- Schouten, R.E., Tijskens, L., van Kooten, O., 2002. Predicting keeping quality of batches of cucumber fruit based on a physiological mechanism. *Postharvest Biol. Technol.* 26 (2), 209–220.
- Schouten, R.E., Zhang, X., Verschoor, J.A., Otma, E.C., Tijskens, L.M.M., van Kooten, O., 2009. Development of colour of broccoli heads as affected by controlled atmosphere storage and temperature. *Postharvest Biol. Technol.* 51 (1), 27–35. <https://doi.org/10.1016/j.postharvbio.2008.06.005>.
- Sun, Z., Yang, J., Hu, D., Tian, H., Ying, Y., Xie, L., 2024. Using knowledge-guided temperature correction for online non-destructive detection of soluble solids content in pear via Vis/NIR spectroscopy. *Postharvest Biol. Technol.* 218, 113178.
- Tijskens, L.M.M., Polderdijk, J.J., 1996. A generic model for keeping quality of vegetable produce during storage and distribution. *Agric. Syst.* 51 (4), 431–452. [https://doi.org/10.1016/0308-521X\(95\)00058-D](https://doi.org/10.1016/0308-521X(95)00058-D).
- Tolesa, G.N., Workneh, T.S., Melesse, S.F., 2018. Logistic regression analysis of marketability of tomato fruit harvested at different maturity stages and subjected to

- disinfection, storage condition and storage period treatments. *Biol. Agric. Hortic.* 34 (1), 40–52.
- Torkashvand, A.M., Ahmadi, A., Nikraves, N.L., 2017. Prediction of kiwifruit firmness using fruit mineral nutrient concentration by artificial neural network (ANN) and multiple linear regressions (MLR). *J. Integr. Agric.* 16 (7), 1634–1644. [https://doi.org/10.1016/S2095-3119\(16\)61546-0](https://doi.org/10.1016/S2095-3119(16)61546-0).
- Villanueva, M., Tenorio, M., Sagardoy, M., Redondo, A., Saco, M., 2005. Physical, chemical, histological and microbiological changes in fresh green asparagus (*Asparagus officinalis*, L.) stored in modified atmosphere packaging. *Food Chem.* 91 (4), 609–619.
- Wang, J., Wang, J., Chen, Z., Han, D., 2017. Development of multi-cultivar models for predicting the soluble solid content and firmness of European pear (*Pyrus communis* L.) using portable vis–NIR spectroscopy. *Postharvest Biol. Technol.* 129, 143–151.
- Wang, N.-N., Sun, D.-W., Yang, Y.-C., Pu, H., Zhu, Z., 2016. Recent advances in the application of hyperspectral imaging for evaluating fruit quality. *Food Anal. Methods* 9, 178–191.
- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., Girshick, R., 2019. Detectron2.
- Xiao, X., Li, M., 2022. Fusion of data-driven model and mechanistic model for kiwifruit flesh firmness prediction. *Comput. Electron. Agric.* 193, 106651.
- Zhu, H., Chu, B., Fan, Y., Tao, X., Yin, W., He, Y., 2017. Hyperspectral imaging for predicting the internal quality of kiwifruits based on variable selection algorithms and chemometric models. *Sci. Rep.* 7 (1), 7845.