Laser scattering imaging combined with CNNs to model the textural variability in a vegetable food tissue

Samuel Verdú [*], José M. Barat, Raúl Grau

Departamento de Tecnología de Alimentos, Universitat Politècnica de València, Spain

A B S T R A C T

In this work, the capacity of the laser scattering imaging, combined with predesigned convolutional neural networks was studied to model the textural variability of pre-processed banana as vegetable tissue. Three different imaging embedders (*Inception V3*, *Painters* and *VGG19*) and four types of images (based on colour channels *RGB*) were tested for that purpose. Two batches of bananas were stored for 0, 3, 7 and 10 days. Fruits were sliced and analysed in imaging and firmness terms. Texture differences were observed for different storage times and fruit zone. Firmness mean and dispersion decreased with storage time. Images from samples revealed light-transmission decreased with storage time. After multivariate statistical procedures, imaging data showed similar behaviour to firmness concerning dispersion. Regression studies evidenced dependency between imaging and firmness data, albeit different errors were obtained depending on the convolutional network and image types. Moreover, after dividing samples into four different models of multilevel categories, successful predictions were obtained from all networks in function of the used image type. Overall, information from the *R* channel improved results for statistical analytics. However, *Inception V3* and *Painters* also present good results from isolated information of *G* and *B*. Results evidenced the capacity of the used imaging procedure to capture and model the variance generated in texture due to tissue differences produced by both storage process and fruit zone.

Laser scattering imaging Convolutional neural networks, Banana, RGB, Texture

1. Introduction

The food industry continuously improves control procedures to ensure the capacity of processes managements and therefore ensure quality and safety standards in the offered products. The main developments focus on substituting outdated devices and techniques for improving control and inspection tasks. These developments reduce energy costs, product heterogeneity, and operation time requirements. That is possible by making applications able to increase the capacity to collect information from the process and food matrix in a given operation. Major analysis velocity means a major number of analysed samples and therefore, higher representativity of quality control data. Adapting new techniques to analyse control procedures and products in different process operations can lead to major short-term improvements without incurring high economic and time costs (Abdul Halim Lim, 2015; Lim, 2016). After data analytics, management could be carried out by databases to improve decision-making about any actuation required in real-time or a posteriori. These approaches are performed mainly by techniques that operate non-destructively (Chen, 2013; Ropodi, 2016). Such techniques need to be based on known physico-chemical phenomena of interest, from which data are collected, but without coming into contact with the product matrix in any measurement.

In this sense, imaging applications represent a vast development area for food processing control improvements. Diverse imaging strategies, such as RGB colour features, texture and shape features, multispectral and hyperspectral imaging, etc., have characterised multiple processes (Caballero et al., 2017; Darnay et al., 2017; Shi et al., 2018). In line with this, the laser scattering imaging technique has been applied to both continuous model processes and characterise end products. It is based on capturing in images the diffraction patterns generated onto the sample surface after transmitting a coherent light (laser) across the food matrix from an opposite surface. Our research group has applied this imaging technique to model features in food processing and products related to textural changes, such as predicting textural properties of vegetable creams (Verdú et al., 2021), modelling the effect of fibre enrichment on biscuit texture (Verdú et al., 2020), modelling texture of milk during the curdling process (Verdú et al., 2021b), among others. In contraposition of another type of imaging techniques, this configuration of laser scattering imaging allows capture information, not only from surface of food but also from internal structures. That property facilitates capturing variability from internal food matrix modifications and therefore enables modelling texture properties non-destructively. In contraposition, diffraction patterns captured in images represent abstract shapes that combine colour patterns apparently random. These abstract shapes require potent feature extraction strategies to analyse the collected variance correctly. That fact is accentuated when a non-continuous matrix is analysed. Different tissues or heterogeneous structures increase the complexity of patterns, requiring deep features extraction to avoid the possible noise provided by natural variability of the studied matrix.

Deep learning technology represents a group of methods that could play a significant role in that requirements. One of the main procedures within computer vision is convolutional neural networks (CNNs). The CNN represents a class of deep feedforward neural networks constructed by imitating the connective pattern of neuronal tissue in the human visual cortex (Hussain et al., 2019). The basic hierarchical structure of CNNs has three main layers: convolution layer, pooling layer, and fully connected layers (among others, Fig. 1). Each transformation layer has a specific role in transforming the input data into a 1D feature vector, which is the input data for applying multivariate statistical approaches. Thus, CNN's can be used as descriptor generators that produce lower dimensionality representations of input instances (automatic feature extractors).

In the last ten years, researchers have done extensive and meaningful research on CNNs focused on detecting and classifying agricultural products. Some examples are used CNN as deep feature extractors to capture the features of diverse food image datasets (McAllister et al., 2018), which were subsequently fed into the multivariate analytics algorithms to perform food tissue classification, classification of damaged lemons based on apparent defects (Jahanbakhshi et al., 2020), inspection and detection of defective maize kernels (Ni et al., 2019), classification of meat origin (Al-Sarayreh et al., 2020), recognition of the browning stages of bread crust (Cotrim et al., 2020), etc. Concerning defects in apple surfaces, achieving a higher recognition rate of 92.5%. Concretely for bananas, some authors reported a successful relationship between texture/ripeness status of bananas using CNNs combined with imaging techniques; however, those studies were principally focused on analysing external colorimetric properties of the whole fruit (Saragih and Emanuel, 2021; Xie et al., 2017; Zhuang et al., 2019).

In this experiment, the banana slices was used as food matrix, where the variability of that tissue in textural terms was modelled using laser scattering and CNNs as rapid and non-destructive analytical method. In this sense, the texture of banana tissue is commonly evaluated by measuring instrumentally parameters such as shear force, firmness, force relaxation, texture profile analysis (TPA), where elasticity, cohesiveness, chewiness and resilience could be calculated (Chauhan et al., 2006). Textural data represent crucial information about the physicochemical status of the tissues and help to make decisions about necessities of storage times, processing conditions and suitability of that raw matter to be derived for transformation procedures in other products.

Therefore, increasing velocity and efficiency in collecting texture data, by techniques such imaging analysis and machine learning procedures, could represent an improvement in quality control for this type of raw matter. In this case, that sliced tissue was selected because the textural properties of bananas are strongly related to the identity properties of processed products from sliced banana such as chips, affecting mainly oil absorption, texture, and sensory score after the frying/drying process (Ammawath et al., 2001). The procedure allowed to know the evolution of the texture homogeneity within the fruit tissues across the storage process. This approach focused towards facilitate a possible optimisation of the selection process of the most suitable slices (in tissue properties terms) in a chips production chain.

Thus, this work was focused on testing the capacity of the laser scattering imaging, combining with CNNs, to model the textural variability in a vegetable food tissue such banana.

2. Material and methods

2.1. Experiment procedure

The experiment was carried out following the scheme included in Fig. 2. The first step was to process the bananas from different batches and storage times to obtain the samples, which were slices of 1 cm thickness. The second step was to capture images of diffraction patterns generated onto samples due to laser-matrix interaction after light transmission. Images were processed selecting the region of interest (ROI) and removing the captured area's rest. After ROI selection, the obtained images were processed to obtain different types of images based on colour channels. All generated images were processed with convolutional neural networks (CNNs) to transform images into features

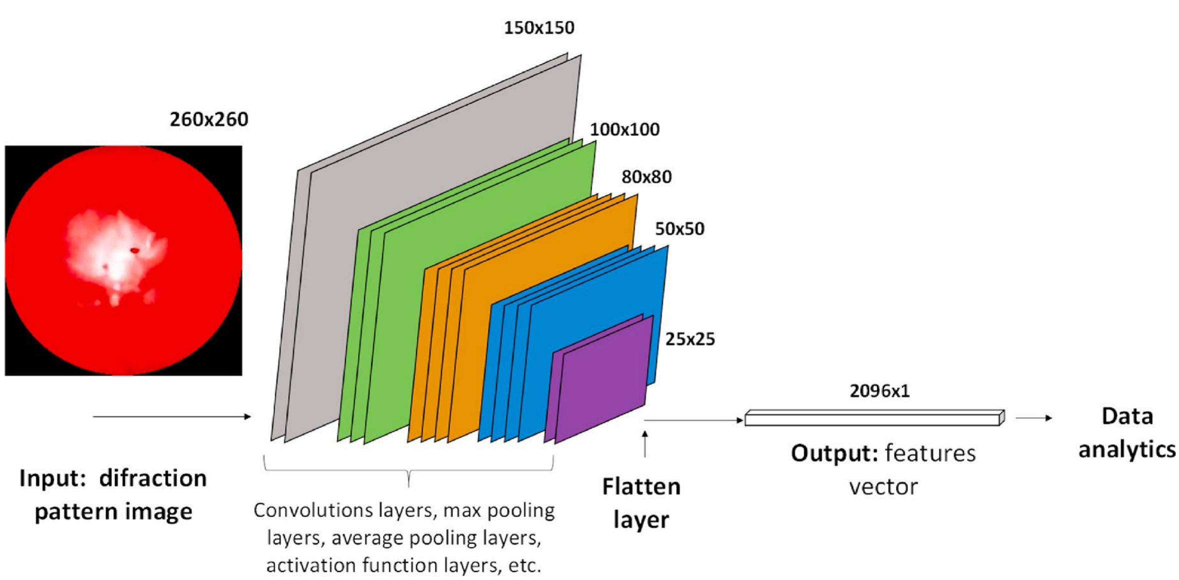

Fig. 1. Scheme of convolutional neural network.

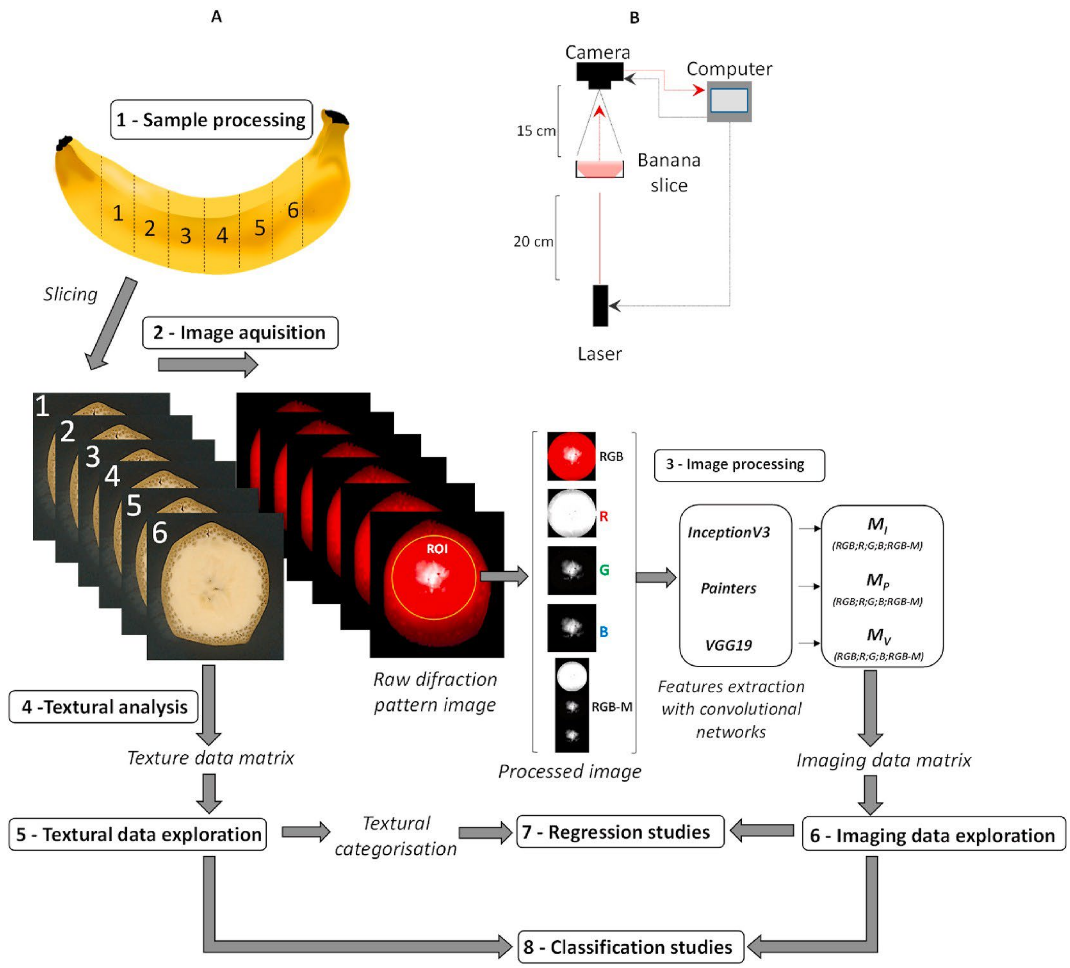

Fig. 2. Scheme of the experimental procedure and imaging device (A).

vectors and thus to create the image data matrix. After imaging procedures, the firmness of all samples was analysed with a texture analyser. Firmness data were organised in a matrix and explored to determine the effect of storage time and the fruit zone (Fig. 2-A, Sample processing, selected zones between the fruit extremes: 1, 2, 3, 4, 5 and 6). In the same way, the impact of those factors on information captured in images and extracted with each CNN was explored. After those singly studies, the relationship between both data sets was tested. It was done in two ways: regression and classification procedures. The regression studies provided information about the dependence of both data sets. In contrast, classification ones offered the capacity of each combination of the type of image and CNN to predict the tissue status classifying samples within different firmness ranges (categories).

2.2. Sample processing

The used vegetable food tissue was banana (*Musa acuminata Colla*) acquired at a local supermarket. In order to obtain a wide range of textures, different storage times from two fruit batches (A and B) of 20 units were studied two weeks apart. The pre-processed fruit dimensions were: 5.3±.5 cm length, 20±2 g weight and 4.3 ± 0.5 cm maximum diameter. The initial ripening stage of samples were between 5 and 6 levels of Von Loesecke and Willard (949) scale concerning peel properties (5 = yellow with green edges; 6 = yellow).

The batch was divided into four groups of five bananas each to study the texture evolution at four different points of storage times (0, 3, 7 and 10 days of storage at 20 °C). At each storage time, samples were obtained from bananas by cutting six transversal slices of 1 cm thickness following the patterns shown in Fig. 2, Sample processing section. Finally, 240 slices were tested (30 samples per storage time and batch).

2.3. Image acquisition

A low-cost imaging device was designed based on Verdú et al. (2020) to collect information from each sample generated by light-slice interactions. The objective was to study images of diffraction patterns, where variance generated in the tissue of product due to storage process and fruit-zone was collected. After laser-matrix interaction, the diffraction patterns were generated onto the sample surface by light transmission. Fig. 2-B shows a scheme of the device setup. The elements were a digital camera, a laser light, and a computer. The setup was installed away from the light inside a dark chamber. A digital Logitech C920 camera (maximum resolution of 1080p/30 fps - 720p/30 fps) was the capture system. Images were saved as. JPEG (1920 × 1080) and were captured in the RGB (red, green and blue) format. The camera was vertically placed 15 cm over the banana slice, which lies in the middle of the visual field. The laser pointer (650 nm, 50 mW) was perpendicularly placed 20 cm under the zone where the slices were placed onto a transparent platform. This type of light was selected because of its coherent properties in wavelength, direction and phase terms. The laser interacted firstly with the endocarp zone, from which the light propagated to the rest of the material. This phenomenon provided representative information from the entire tissue.

A computer controlled both the camera and the laser line. Images were manually captured by placing the slices individually into the imaging device.

2.4. Image processing

Information collected in images was extracted after processing raw images. ROIs (region of interest) from each image was isolated by cutting raw images in the central zone of the diffraction pattern (Fig. 2) using a circular area of 260 pixels diameter. That dimension was selected as the area that collected the maximum diffraction pattern information for all slices, independently of sample size. Thus, the ROIs were processed to obtain R, G and B versions of the images by splitting the colour channels. Moreover, considering the high capacity of CNNs to extract information from highly complex images, we decided to include a mode of image where the information from diffraction patterns captured by each channel was included in the same 2D grayscale image with spatial independence. These images were generated by concatenating *R*, *G* and *B* images into one, called *RGB-M* (Fig. 2, Processed image). Therefore, five types of images were studied in this experiment: *RGB, R, G, B* and *RGB-M*. The information collected from G and B channels was studied because common sensors tend to isolate the R, G and B bands of the visible light spectrum as they are seen by human eyes, but some zones overlapped because laser excites sensors G and B, albeit less efficiently (Batistell et al., 2014). These zones could offer extra data, which can be useful for this type of imaging procedures and analysis.

Information from the processed images was extracted using three different CNN as feature generators. The objective was to capture deep information from diffraction patterns throughout the extracted features, which was the input of posterior multivariate statistical analyses. CNN's were applied by using the imaging analytics module of Orange software (Godec et al., 2019). CNN's are based on computational processing blocks/layers from which images are embedded and transformed to an analysable data vector. This analytic module facilitated the embedding of images by several CNN, pre-designed to process images under different protocols following the main objective of each, although suitable to extract helpful information from different types of objects. The output was a feature vector of each image, whose number of features (variables) depended on the used model. The CNN embedders tested in this case were: *InceptionV3*, *Painters* and *VGG19*. *InceptionV3* is Google's deep neural network, trained on the ImageNet data set for general objects recognition in digital images. It is a commonly used image recognition model with an accuracy rate greater than 78.1% on the ImageNet dataset (Szegedy et al., 2016). It was selected because of its power and unspecific objective. *Inception-v3* is composed of a 42-layer neural network and consists of symmetric and asymmetric building blocks (Demir et al., 2019). Otherwise, *Painters* is a CNN created by Nejc Ilenic, who won *Kaggle's Painter by Numbers*; a competition challenged to examine paintings from different artists. The CNN was created to inspect pairs of paintings and determine whether they were created by the same artist. It is a CNN embedder based on a 24-layer neural network, trained on 79,433 images of paintings by 1584 different painters in the training set (https://github.com/inejc/painters). It was selected based on the abstract features of images from diffraction patterns and its similarity to several painting styles. The third CNN used was *VGG19.* It is a deep neural network for image analysis and recognition created by the Visual Geometry Group from the University of Oxford, consisting of 19-layers. It is trained on the ImageNet data set and performs 7.3% top-5 errors at the ImageNet dataset (Godec et al., 2019). All of the mentioned models are based on the same function blocks, differently ordered depending on the architecture pre-designed by creators, including convolutions layers, max-pooling layers, average pooling layers, activation function layers, flatten layer, etc. (Gu et al., 2018). Therefore, embedders were used as powerful features generators that produced lower dimensioned representations from each image. The last layer was not used to avoid the pre-established conditions of each CNN for the modelling process. Based on this utility, the used information was the output from the activation layers of the penultimate layer of the models, which generated feature data from the images previously to the training/prediction process of CNN itself. Thus, data matrix from diffraction patterns of banana-slices samples were generated based on the three above mentioned embedders. The number of features generated per image from each CNN was 2048 for *InceptionV3* and *Painters* and 4096 for *VGG19*.

2.5. Textural analysis

After imaging capture, variability associated with tissues changes were studied by texture analysis. The used parameter was firmness, which was taken using the texture analyser (TA-XT2i, Stable Microsystems Texture Technologies Inc., UK). Comprehension assays were carried out on banana slices using a 30 mm diameter flat probe, which covered totally the endocarp and the most of mesocarp, in order to obtain representative textural information from entire tissue. The test speed was 1 mm/s, and the penetration depth was 50% thickness (5 mm) (Wang et al., 2013). The firmness unit was expressed as N.

2.6. Statistical analysis

The texture results were analysed by ANOVA to evaluate differences in fruit batches across storage time. In those cases in which the effect was significant (P-value < 0.05), means were compared by Fisher's least significant difference (LSD) procedure. The image information was explored by applying multivariable statistical procedures to reduce dataset dimensionality and evaluate its dependency on texture data. To this end, the multivariate unsupervised statistical method Principal Component Analysis (PCA) was used to explore the structure of data variance extracted by CNNs from the different types of images. Moreover, the multivariate supervised method Partial Least Square Regression (PLS-R) was used to evaluate the dependency between the image information and the texture data as a continuous variable. This method was used to carry out the linear regression models between both datasets, which were evaluated based on the R^2 of calibration and cross-validation (venetian blinds) and root mean square error (RMSE), which was calculated following equation (1):

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y'_i - y_i)^2}{n}} \tag{1}$$

were n being the number of samples, y'_i is the mean value and y_i a value from a given sample.

After regression studies, the capacity of the technique to classify the samples in discrete texture ranges was explored. In this study, samples were divided into 2, 3, 4 or 5 ranges of texture to generate classification models with increasing complexity. A support vector machines-discriminant analysis (*SVM-DA*) was used to study the capability of the imaging technique to predict the range of texture of samples within these four different models of categorial subdivision. Prediction studies were done by establishing randomised samples as training batches (66%) to generate the models and the remaining samples as the test lots (33%) to be predicted. The results were expressed in precision and ac- curacy terms based on the following equations:

$$accuracy = \frac{TN + TP}{TP + TN + FP + FN} \tag{2}$$

$$precision = \frac{TP}{TP + FP} \tag{3}$$

were TN = true negative results, TP = true positive, FP = false positives and FN false negatives. These procedures were run with the PLS Toolbox, 6.3 (Eigenvector Research Inc., Wenatchee, Washington, USA), a toolbox extension in the Matlab 7.6 computational environment (The Mathworks, Natick, Massachusetts, USA).

3. Results and discussion

3.1. Textural data exploration

Texture data from the storage process are included in Fig. 3. A reduction of firmness was observed during storage time for both analysed batches. Both batches had a similar firmness reduction rate, although differences within sampling points were observed. Results from variance analysis showed significant differences among all analysed sample distributions. Tissues from both batches started from different average firmness (t_0) and evolved, showing different firmness levels at each sampling point. Moreover, the evolution pattern of dispersion was observed across storage time. Reduction of dispersion was observed within each data distribution following days of storage. That meant homogenisation of firmness among different fruits and internal zones via natural modifications in the tissues. Texture behaviour observed in results fitted with the extensive literature reporting about the characterisation of banana texture during the ripening process. Compositional changes are associated with biochemical and physiological modifications resulting in softening (Sanaeifar et al., 2016). Concretely, changes in the cell wall structure by enzymes and starch degradation are the two main reasons for fruit softening (Osorio and Fernie, 2013). Cell wall enzyme activities increase the solubility of pectins and hemicelluloses, possibly through a debranching process that decreases polymer interactions (Do Prado et al., 2016). Homogenisation of firmness was a typical fact reported previously in other studies such as Xie et al. (2018), where ten-times-reduction in standard deviation was observed between unripe and ripe bananas. Therefore, regardless of initial firmness level, two differenced effects were evidenced on texture data, storage time and fruit zone, principally for sampling points near t_0. Those effects generated wide variability in samples (slices), representing the variability managed in a real processing chain with this type of transformation in banana fruits. It allowed us to study the response of imaging technique in an important texture range for this type of fruit matrix.

3.2. Imaging data exploration

To preview the effect of texture on the diffraction patterns captured in each type of processed image, samples from different textural levels were explored. Fig. 4 shows images of diffraction patterns generated in samples of the extremes of the obtained texture range, in order to

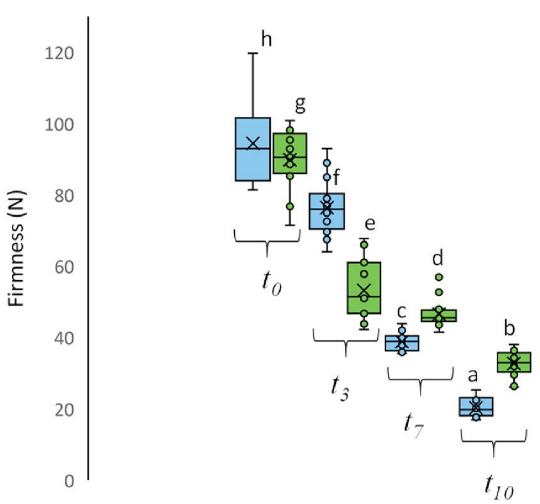

Fig. 3. Texture results. Box and whisker plot of firmness data. Blue series: batch A; green series: batch B. t_n indicates sampling point during storage time in days. Letters denote significant differences at the 95% confidence between data distributions from different sampling points.

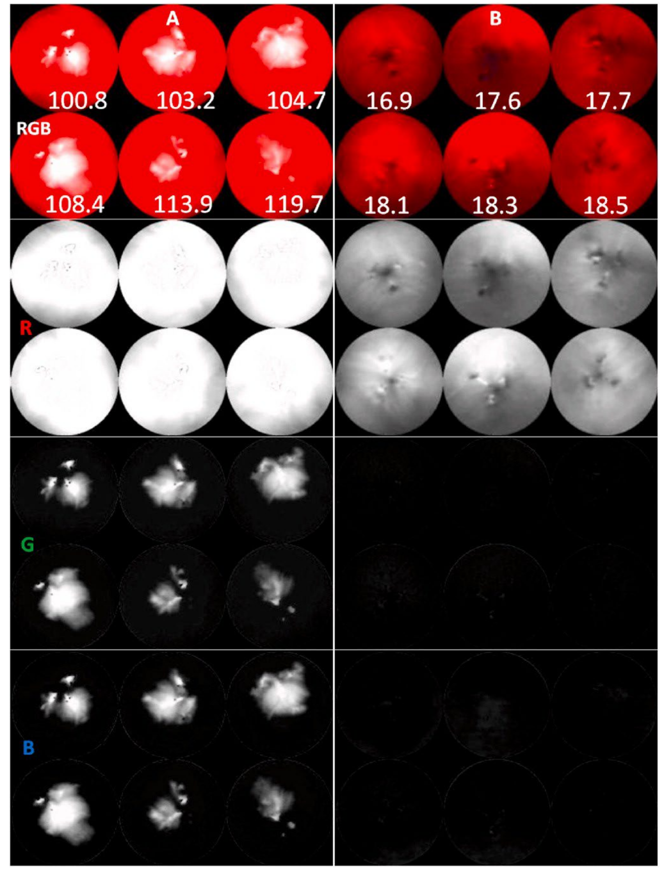

Fig. 4. Real images of diffraction patterns. A: samples from the maximum firmness extreme (distribution h, Fig. 3); B: samples from the minimum firmness extreme (distribution a, Fig. 3). *RGB*, *R*, *G* and *B* indicate the type of image based on colour channels. Numbers indicate the firmness of the sample expressed in *N*.

visualize possible differences between them. Images type *RGB*, and each splinted channel were included. The main differences were observed in colour intensity distribution within the patterns. Samples from the highest extreme (distribution "h" from texture data analysis in Figs. 3, Fig. 4-A) trended to present an important fraction of pixels capturing high intensity (central white zones). In contrast, samples from the lowest extreme (distribution "a" from texture data analysis in Figs. 3, Fig. 4-B) presented lower intensity. That difference was better observable in split channels. Pixels with high intensity formed patterns even in *G* and *B* channels. That effect was produced because sensors from those channels could also be stimulated at high light expositions from different wavelengths (in this case, within the red zone of the spectrum). This phenomenon could be used to capture valuable information lost from the R channel due to the saturation of those zones for the red channel sensor (Grau et al., 2021). Thus, the observed differences in the rate of pixels capturing high intensity indicated modification of light transmittance across the matrix between observed samples. This phenomenon suggested a relationship between the transmittance of banana tissue and texture changes. Moreover, differences in patterns from the softest samples (Fig. 4-B) seemed less than the observed for the samples with high firmness (Fig. 4-A), principally for *G* and *B* channels. Thus, changes produced in banana tissues during ripeness could be fitted with the observed modifications in light transmittance and patterns heterogeneity. The occurred compositional modifications based on the degradation of starch, pectins, and hemicelluloses impacted cell wall structures, altering light transmission by modifying the chemical and physical properties of tissue, respectively (Kulkarni et al., 2011). Thus, the increase of soluble solutes and the collapse of tissue structures could explain decreased light transmittance and heterogenicity reduction of diffraction pattern properties.

After exploring the raw images, the data matrix extracted from each image type, after CNN's embedding, were explored through PCAs. This procedure was carried out to study if imaging technique, combined with CNNs, captured the trend observed in the structure of variance in samples. This non-supervised procedure projected the variance captured in the imaging features to a reduced number of dimensions, in this case, in one called PC1. The information collected in the PC1s represents the most of the variance generated by the features obtained from CNN processing. The aim was to explore easily, without supervision, how the variability of the features evolved with the tissue status. Fig. 5 shows the results of the dimensionality reduction for all CNNs features extracted from the *RGB* images. Samples were ordered following the observed differences in the firmness study. The studied scores were collected in PC1 since significant parallelism in variance structure to that observed in the textural analysis was observed. When samples were ordered following decreasing firmness values, a common pattern was observed for all CNNs but not for all image types. A decreasing trend of PC1 scores was observed for *RGB* images of Painters and *VGG19* and not for Inception V3, fitting with the firmness decreasing (Fig. 5). Dispersion starts from distribution *h*, from which decreasing trend was produced. On the other hand, the observed evolution of dispersion in texture seemed to be captured in score data.

This behaviour was also observed for other image types as well as among CNNs. Fig. 6 shows, as an example, the evolution of scores for *R* and *G* images for all CNNs. *Inception V3* improved in those cases, reporting the abovementioned reduction of variance following firmness decrease. The observed differences for *Inception V3* revealed the importance of the type of image for that CNN, although the results could be different under a supervised analysis method. *Painters* maintained the tendency for *R* and *G* while *VGG19* lost that observable trend, showing no patterns for variance across the distributions of samples. These studies revealed the capacity of CNNs to extract useful information from the changes produced in diffraction patterns due to the modifications in tissue-light interaction during storage. However, differences in the extracted information were observed depending on the used image type and CNNs. Note that PCA tried to explain the maximum variance between samples without considering another variable such as single textures or other categories within samples. It meant that supervised analytics were necessary to carry out to better explore data structure and then measure its dependence on texture data.

3.3. Regression studies

After observing the capacity of CNNs to extract information from diffraction patterns and their parallelism with texture data structure, the dependence of texture and imaging dataset from each CNN and image type were explored by regression studies. The PLS method was used to study the relationship between both data sets, independently of sampling point or fruit zone. The dependency of both data sets was evaluated by R^2 and RMSE of calibration and cross-validation. Table 1 shows the results. *Inception-V3* and *Painters* provided the highest R^2 for both calibration and cross-validation, up to 0.90 and 0.80 respectively for all image types excepting *G* for *Inception-V3*. Otherwise, *VGG19* data had the least capacity to correlate with firmness data. In this case, considerable differences among image types were observed. Concretely, regression from *R* data improved the coefficient for both calibration and cross-validation. In addition, regression for *RGB-M* data had values up to 0.80.

Results revealed the capacity of the used CNNs for extracting variance captured in diffraction patterns due to the modification of laser-banana matrix interaction across the storage process. Thus, depen- dence between imaging data and texture modification was evidenced, although there were important differences concerning to type of embedder and image. Generally, although *VGG19* extracts a higher de- scriptors number than the other used embedders, probably the smaller number of layers ended in less capacity to capture details from variance generated in these images due to texture modifications. Moreover, *G* and *B* data seemed to produce noise effects for *VGG19*. That effect was evi- denced with the improved regression coefficients reported when iso- lated *R* data from images were used. This phenomenon was attributable to the fact that the used laser is formed by 650 nm light, which should excite *R*-channel mostly. Generally, digital cameras do not capture colour in a specific channel as a particular wavelength of the visible spectrum. Typical sensors capture the *R*, *G* and *B* bands of the visible light simulating human eyes, but some adjoining zones are captured in more than one channel, evidencing that the laser also excites sensors *G* and *B*, albeit less efficiently. Thus, information from *G* and *B* channels performed like noise for the *VGG19* case, improving results when it was removed. Note that G and B information could be useful as supplementary information to R-channel in other cases. In the case of *Inception-V3* and *Painters,* extracted information from *G* and *B* modelling texture by themselves. That hypothesis fitted with the also improved result for *RGB-M* for *VGG19*. In that case, data from each channel were included but isolated in the same layer. Then, the improvement of the results for this type of image was probably produced by the presence of the isolated *R* data, although lower than the case of *R* images because of the presence of *G* and *B* data simultaneously. It indicated that the structure of *RGB* information used as input for CNN could report different results depending on the image format.

To visualize the relationship between imaging data extracted with embedders and firmness observed in the results mentioned above, scores from the first latent variable from PLS (LV1, containing the most proportion of explained variance), used as dimensionally reduced imaging data, was represented vs firmness data. Fig. 7-A and 7-B shows regression plots for *RGB* and *R* types of images from *VGG19,* respectively. As observed in Table 1, *VGG19* presented disrupted linearity for *RGB*;

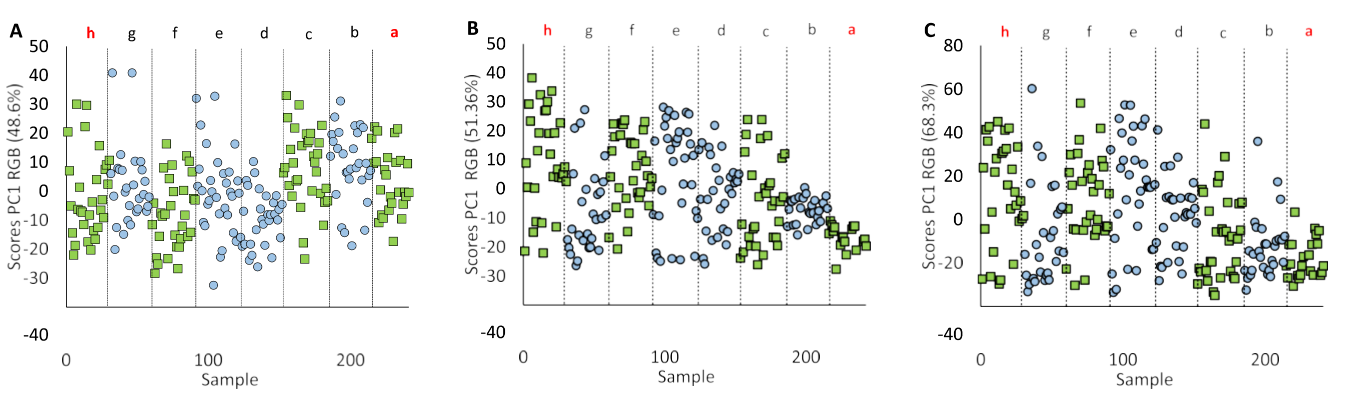

Fig. 5. PCA studies from *RGB* images data extracted with CNNs. Scores of PC1 from A: *Inception-V3*; B: *Painters*; C: *VGG19*. Blue series ●: batch A; green series ■: batch B. Letters indicate differences from ANOVA study on texture data (Fig. 3). Red letters indicate groups that belong to samples shown in Fig. 4.

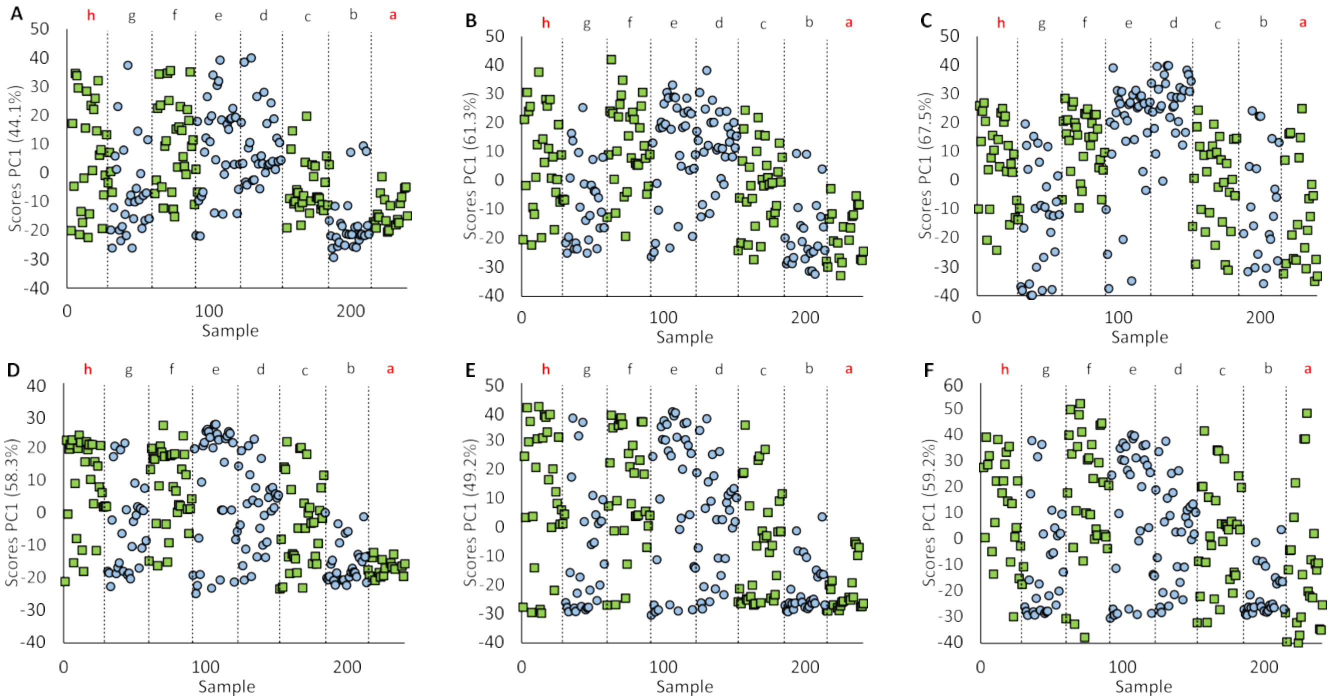

Fig. 6. Examples of PCA studies from R and G images data extracted with CNNs. Scores of PC1 from A: *R* x *Inception-V3*; B: *R* x *Painters*; C: *R* x *VGG19*; D: *G* x *Inception-V3*; E: *G* x *Painters*; F: *G* x *VGG19*. Blue series ●: batch A; green series ■: batch B. Letters indicate differences from ANOVA study on texture data (Fig. 3). Red letters indicate groups that belong to samples shown in Fig. 4.

Table 1
Results of regression studies.

		RGB	R	G	B	RGB-M
Inception-V3	***R²cal***	**0.99**	**0.97**	**0.99**	**0.99**	**0.97**
	R²cv	**0.85**	**0.83**	0.77	**0.85**	**0.81**
	RMSEcal	1.42	5.15	2.02	0.96	5.13
	RMSEcv	9.17	9.73	12.15	9.25	11.20
Painters	***R²cal***	**0.99**	**0.99**	**1.00**	**0.99**	**1.00**
	R²cv	**0.86**	**0.84**	**0.86**	**0.82**	**0.86**
	RMSEcal	1.31	0.98	0.40	0.55	0.62
	RMSEcv	9.17	9.27	8.98	9.76	9.01
VGG19	***R²cal***	0.73	**1.00**	0.79	0.69	**0.89**
	R²cv	0.46	**0.82**	0.27	0.35	**0.83**
	RMSEcal	12.03	0.52	11.47	13.82	7.48
	RMSEcv	18.60	10.20	23.98	22.30	10.19

cal: calibration; *cv*: cross-validation; *RMSE*: root mean square error.

however, high linearity was obtained when R data was isolated.

Other imaging colour features (L*, a*, b*, H, S, V, Y, U, and V) and hyperspectral imaging have been used to model banana firmness with different ripening levels, obtaining high correlations between imaging and firmness data. Cho and Koseki (2021) obtained coefficients R^2 = 0.97 between colour features from banana images and firmness using neural networks. Moreover, Xie et al. (2017) predicted firmness in bananas (R^2 = 0.74) after exploring and isolating the relevant wavelengths from hyperspectral information.

The main differences with those approaches fall on that, in those cases, imaging information was based on colour features from whole banana skins and average of firmness, while here, the characterisation was focused on each banana slice. That type of strategy based on external colour features could also be useful to measure some properties of banana slices. However, the contact with oxygen after the slicing process would produce browning quickly, which should not be related to previous ripening and then generate variability not associated with structural tissue changes. In this sense, diffraction patterns collected information from inside the sample after light transmission across the banana matrix, which seemed to capture variance provided by fruit zones and the storage effect.

3.4. Classification studies

After evidence of the dependence between imaging data and texture from banana samples, the total number of samples (slices), from day 0–10, were divided into a different number of categories (2, 3, 4 and 5). Variance analysis showed eight samples distributions with significant differences in their average texture (Fig. 3), however there were overlapped zones where slices from different groups had similar values (Fig. 1-Supplementary material). Then, instead of using those textural groups, new specific limits on the raw texture dataset were established to obtain more accurate thresholds (Fig. 2-Supplementary material). Therefore, the texture categories were established independently of the storage time, batch and fruit zone, in order to focus the division exclu- sively on the variability observed for texture status. It was carried out by segmenting the range of texture from the obtained results in equative segments depending on the number of categories. That is to say; all samples are in 2 categories, all in 3, all in 4 or 5. The aim was to evaluate and quantify the ability of the imaging technique and CNNs to predict and classify in different textural levels. Results are included in Fig. 8 in accuracy and precision terms. Overall, values of classification parameters were reduced, following the number of categories increasing. That tendency was due to the rising complexity of the model. Moreover, differences among CNNs and image types were observed. *Painters* presented the best results with all combinations (image types x number of categories) up to 0.8 for both precision and accuracy, except *B*-4 and *B*-5 (Fig. 8-B). Concretely, all combinations for *R* provide coefficients near to 1, therefore improving the results of data from all channels such as *RGB* and *RGB-M*, principally for models with five categories. *Inception-V3* also had high coefficients for most combinations; however, many of them did not rise 0.9 in precision or 0.8 in accuracy (Fig. 8-A). In this case, *R* data also improved the prediction of texture categories, principally until four

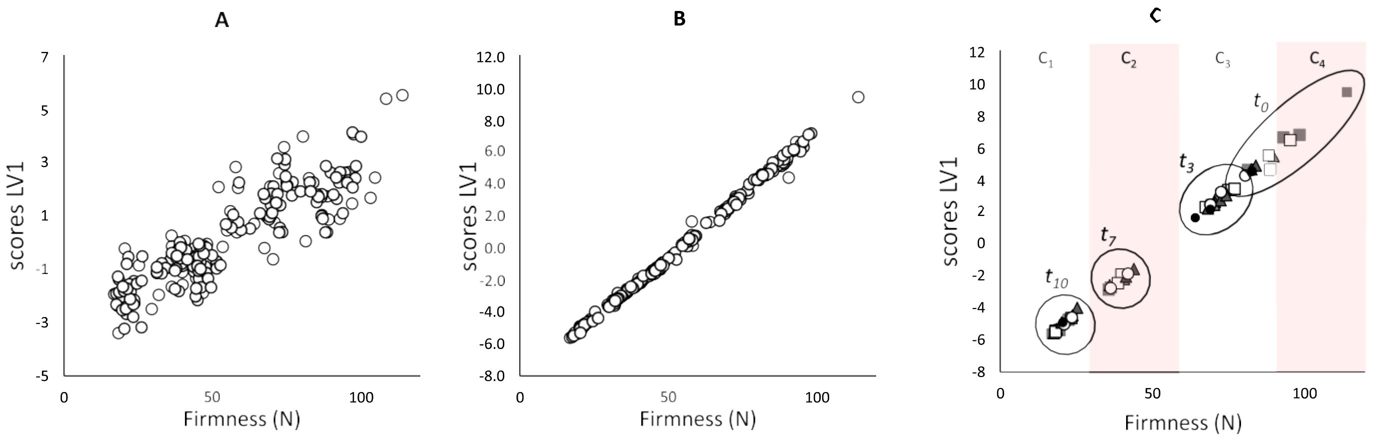

Fig. 7. Regression results between different imaging processed data after *PLS* modelling and firmness. A: *LV1* scores from *RGB* images data extracted with *VGG19* vs firmness; B: *LV1* scores from *R* images data extracted with *VGG19* vs firmness; C: *LV1* scores of batch A from *R* images data extracted with *VGG19* vs firmness. Series in C indicates different fruit zones 1: ■; 2: □; 3: ▲ ; 4:▲ ; 5:○ ; 6:● t_n indicates sampling point during storage time in days. Watermarks indicate the four firmness categories (C_1, C_2, C_3 and C_4).

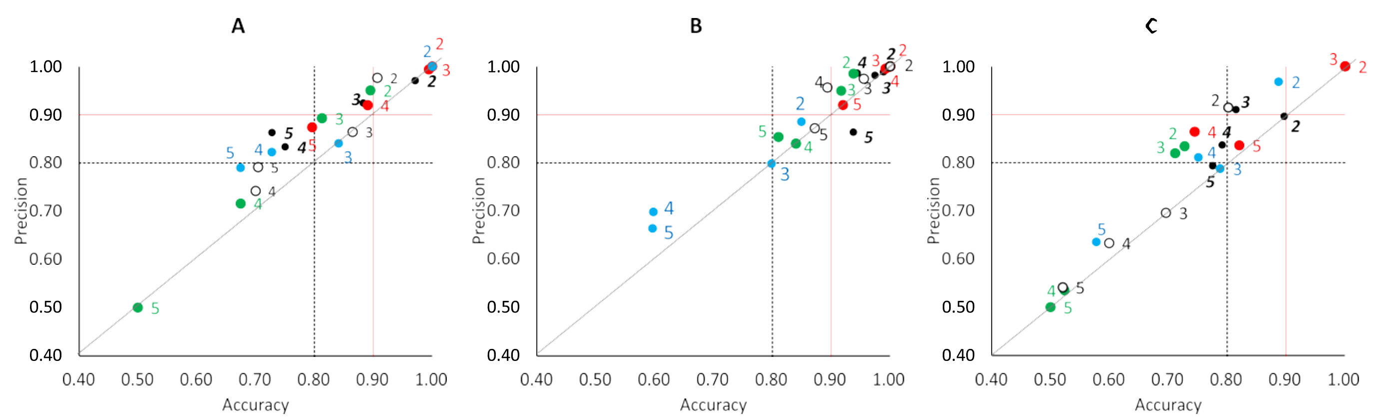

Fig. 8. Results of classification studies. A: Inception-V3; B: Painters; C: VGG19. Colour indicates the image type: R: ●G: ●B: ●RGB: ●; RGB-M: ○. Numbers indicate the number of categories for this result. The dotted line indicates a threshold of 0.80 and the continuous red line indicates a threshold of 0.90 for both precision and accuracy.

levels. *Inception-V3* showed displacement trending to high precision values but reduced accuracy simultaneously. *VGG19* had the highest number of combinations under 0.8 (Fig. 8-B). *G*, *B* and *RGB-M* presented the worst results. In this case, *R* raised nearly 1 for models with 2 and 3 categories. It was according to the above-observed improvement of regression coefficients for *R* data. *VGG19* also showed displacement trending to high precision values but reduced accuracy. Thus, *Painters* was the unique embedder supporting high classification parameters (up to 0.80) for five different categories.

Note that the fact that regression studies showed better adjust than the classification ones could be explained because the categorisation of textures was established based on non-natural thresholded between samples. The imaging information depended on the textural tissue state, a continuous phenomenon. That means the existence of samples placed at the extreme of a category near other samples placed in the extreme of the contiguous category could increase the probability of classification errors.

Thus, results showed a similar capacity of CNNs to differentiate bananas from a different ripening process using imaging techniques focused on analysing external colorimetric properties of the whole fruit. Several examples are Saragih and Emanuel (2021), who evaluated pre-trained CNN models to classify four different ripeness levels of the banana fruit, whose accuracy was up to 0.90. Xie et al. (2017), successfully classified bananas in a binary model riped-unripped, however in this case, hyperspectral imaging data was used. Moreover, Zhuang et al. (2019) used colour, local texture and shape features from external properties of fruit classifying bananas in four ripening levels. These types of features were also used by Mazen and Nashat (2019) for classifying bananas in a model with four ripeness levels, but in this case, an artificial neural network to process imaging data was used.

Results evidenced the capacity of the technique for extracting enough information from light-matrix interaction to predict the texture range of samples in models with varying complexity. It meant texture variance in this type of food matrix was successfully modelled as the unique dependent variable, despite the cause of variation being fruit zone or storage time. Thus, considering the increasing homogeneity in texture with storage process, fruits from early storage times should provide slices to different categories concerning the above-established ones.

To evidence that hypothesis, regression data of batch A from *R* images data extracted with *VGG19* (Fig. 7-B) were isolated and plotted alone in Fig. 7-C, labelling the sliced samples in series following the six different positions of banana slices (Fig. 2). The firmness axis (X) has been divided (as an example) into the four categories defined for the classification studies. Results showed how samples from bananas at t_0 had slices belonging to two contiguous categories (C_3 and C_4), while the rest of the groups were placed within a single category. Those differences should be done due to the fruit-zone effect, according to the decreasing dispersion tendency following storage time degree observed in both textural analysis (Fig. 3) and imaging data exploration with PCA (Figs. 5 and 6). Thus, fruits from a group with a low storage time (high firmness) should present a higher distance between samples. That is to say, samples from different fruit zones were placed in different categories simultaneously (categories established in classification studies), while high storage time bananas kept all slices within a single group.

Thus, results revealed the capacity of the imaging technique and CNNs to extract enough information from the interaction of laser-matrix to predict the range of firmness of banana slices within multilevel classification models. This property was highly influenced by both the type of CNNs and the type of image used as input.

4. Conclusions

Results evidenced the capacity of the imaging technique and CNNs to capture and model the variance generated in the texture of the banana tissue produced by both the storage time and fruit zone. Alterations produced in banana tissues generated variability in firmness, also modifying light-matrix interactions. Those modifications generated variance within diffraction patterns, which could be captured by the used imaging system. At the same time, the information contained in that captured variance was successfully extracted by processing images with different CNNs. Moreover, information collected in captures depended on the type of analysed image. Features extracted by Inception V3 and Painters modelled texture variability successfully from any type of image under supervised analysis such as the regression studies, while VGG19 seemed to need isolated information from the R channel. Therefore, data extraction could be optimised exclusively to the R im- ages for the three used CNN. The proved dependence between extracted imaging data and texture variance made it possible to predict the range of textures, although in this case, an increasing number of categories produced a reduction of accuracy and precision for all CNNs. The results revealed that, although the CNNs were optimised to model specific types of objects, they could also be used to model abstract structures such as the ones studied in this experiment. This application could be integrated for fast characterisation of banana samples in the production chain to improve the subsequent transformation processes and then to ensure higher homogeneity of texture features of the end product such as “ba- nana chips”. The model could be enriched by including images from longitudinal slices of banana tissue or from a portion not considering the endocarp in future experiments. The next step will be to design specific CNNs for this type of application, which focus to maximise the extraction of texture-dependent information from diffraction patterns of this type of vegetable tissues.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors gratefully acknowledge the financial support from the University Polytechnic of Valencia.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jfoodeng.2022.111199.

References

Abdul Halim Lim, S., 2015. Towards a conceptual roadmap for statistical process control implementation in the food industry. Trends Food Sci. Technol. https://doi.org/10.1016/j.tifs.2015.03.002.

Al-Sarayreh, M., Reis, M.M., Yan, W.Q., Klette, R., 2020. Potential of deep learning and snapshot hyperspectral imaging for classification of species in meat. Food Control. https://doi.org/10.1016/j.foodcont.2020.107332.

Ammawath, W., Che Man, Y.B., Yusof, S., Rahman, R.A., 2001. Effects of variety and stage of fruit ripeness on the physicochemical and sensory characteristics of deep-fat-fried banana chips. J. Sci. Food Agric. 81, 1166–1171. https://doi.org/10.1002/jsfa.922.

Batistell, G., Zhang, V.C., Sturm, J., 2014. Color Recognition Sensor in Standard CMOS Technology. Solid. State. Electron. https://doi.org/10.1016/j.sse.2014.06.002.

Caballero, D., P´erez-Palacios, T., Caro, A., Amigo, J.M., Dahl, A.B., ErsbØll, B.K., Antequera, T., 2017. Prediction of pork quality parameters by applying fractals and data mining on MRI. Food Res. Int. https://doi.org/10.1016/j.foodres.2017.06.048.

Chauhan, O.P., Raju, P.S., Dasgupta, D.K., Bawa, A.S., 2006. Instrumental textural changes in banana (var. Pachbale) during ripening under active and passive modified atmosphere. Int. J. Food Prop. https://doi.org/10.1080/10942910600596282.

Chen, Q., 2013. Recent advances in emerging imaging techniques for non-destructive detection of food quality and safety. Trends Anal. Chem. https://doi.org/10.1016/j.trac.2013.09.007.

Cho, B.H., Koseki, S., 2021. Determination of banana quality indices during the ripening process at different temperatures using smartphone images and an artificial neural network. Sci. Hortic. (Amst.). https://doi.org/10.1016/j.scienta.2021.110382.

Cotrim, W. da S., Minim, V.P.R., Felix, L.B., Minim, L.A., 2020. Short convolutional neural networks applied to the recognition of the browning stages of bread crust. J. Food Eng. https://doi.org/10.1016/j.jfoodeng.2020.109916.

Darnay, L., Kr´alik, F., Oros, G., Koncz, A´., Firtha, F., 2017. Monitoring the effect of transglutaminase in semi-hard cheese during ripening by hyperspectral imaging. J. Food Eng. https://doi.org/10.1016/j.jfoodeng.2016.10.020.

Demir, A., Yilmaz, F., Kose, O., 2019. Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3. In: TIPTEKNO 2019 - Tip Teknolojileri Kongresi. https://doi.org/10.1109/TIPTEKNO47231.2019.8972045.

Do Prado, S.B.R., Melfi, P.R., Castro-Alves, V.C., Broetto, S.G., Araújo, E.S., Do Nascimento, J.R.O., Fabi, J.P., 2016. Physiological degradation of pectin in papaya cell walls: release of long chains galacturonans derived from insoluble fractions during postharvest fruit ripening. Front. Plant Sci. https://doi.org/10.3389/fpls.2016.01120.

Godec, P., Panˇcur, M., Ileniˇc, N., Cˇopar, A., Straˇzar, M., Erjavec, A., Pretnar, A., Demˇsar, J., Stariˇc, A., Toplak, M., Zˇagar, L., Hartman, J., Wang, H., Bellazzi, R., Petroviˇc, U., Garagna, S., Zuccotti, M., Park, D., Shaulsky, G., Zupan, B., 2019. Democratized image analytics by visual programming through integration of deep models and small-scale machine learning. Nat. Commun. https://doi.org/10.1038/s41467-019-12397-x.

Grau, R., Verdú, S., P´erez, A.J., Barat, J.M., Talens, P., 2021. Laser-backscattering imaging for characterizing pork loin tenderness. Effect of pre-treatment with enzyme and cooking. J. Food Eng. https://doi.org/10.1016/j.jfoodeng.2021.110508.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent Advances in Convolutional Neural Networks. Pattern Recognit. https://doi.org/10.1016/j.patcog.2017.10.013.

Hussain, G., Maheshwari, M.K., Memon, M.L., Jabbar, M.S., Javed, K., 2019. A CNN Based Automated Activity and Food Recognition Using Wearable Sensor for Preventive Healthcare. Electronics. https://doi.org/10.3390/electronics8121425.

Jahanbakhshi, A., Momeny, M., Mahmoudi, M., Zhang, Y.D., 2020. Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks. Sci. Hortic. (Amst.). https://doi.org/10.1016/j.scienta.2019.109133.

Kulkarni, S.G., Kudachikar, V.B., Keshava Prakash, M.N., 2011. Studies on physico-chemical changes during artificial ripening of banana (Musa sp) variety “Robusta. J. Food Sci. Technol. https://doi.org/10.1007/s13197-010-0133-y.

Lim, S.A.H, 2016. Statistical process control readiness in the food industry: development of a self-assessment tool. Trends Food Sci. Technol. https://doi.org/10.1016/j.tifs.2016.10.025.

Mazen, F.M.A., Nashat, A.A., 2019. Ripeness classification of bananas using an artificial neural network. Arabian J. Sci. Eng. https://doi.org/10.1007/s13369-018-03695-5.

McAllister, P., Zheng, H., Bond, R., Moorhead, A., 2018. Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. Comput. Biol. Med. https://doi.org/10.1016/j.compbiomed.2018.02.008.

Ni, C., Wang, D., Vinson, R., Holmes, M., Tao, Y., 2019. Automatic inspection machine for maize kernels based on deep convolutional neural networks. Biosyst. Eng. https://doi.org/10.1016/j.biosystemseng.2018.11.010.

Osorio, S., Fernie, A.R., 2013. Biochemistry of fruit ripening. In: The Molecular Biology and Biochemistry of Fruit Ripening. https://doi.org/10.1002/9781118593714.ch1.

Ropodi, I., 2016. Data mining derived from food analyses using non-invasive/non-destructive analytical techniques; determination of food authenticity, quality & safety in tandem with computer science disciplines. Trends Food Sci. Technol. https://doi.org/10.1016/j.tifs.2016.01.011.

Sanaeifar, A., Bakhshipour, A., De La Guardia, M., 2016. Prediction of banana quality indices from color features using support vector regression. Talanta. https://doi.org/10.1016/j.talanta.2015.10.073.

Saragih, R.E., Emanuel, A.W.R., 2021. Banana ripeness classification based on deep learning using convolutional neural network. In: 3rd 2021 East Indonesia Conference

on Computer and Information Technology. https://doi.org/10.1109/EIConCIT50028.2021.9431928. EIConCIT 2021.
Shi, C., Qian, J., Han, S., Fan, B., Yang, X., Wu, X., 2018. Developing a machine vision system for simultaneous prediction of freshness indicators based on tilapia (Oreochromis niloticus) pupil and gill color during storage at 4 °C. Food Chem. 243, 134–140. https://doi.org/10.1016/j.foodchem.2017.09.047.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2016.308.
Verdú, S., Ruiz-Rico, M., P´erez, A.J., Barat, J.M., Grau, R., 2020. Application of laser backscattering imaging for the physico-chemical characterisation of antimicrobial silica particles functionalised with plant essential oils. J. Food Eng. https://doi.org/10.1016/j.jfoodeng.2020.109990.
Verdú, S., P´erez, A.J., Barat, J.M., Grau, R., 2021. Non-destructive control in cheese processing: modelling texture evolution in the milk curdling phase by laser backscattering imaging. Food Control. https://doi.org/10.1016/j.foodcont.2020.107638.
Wang, Y., Luo, Z., Du, R., Liu, Y., Ying, T., Mao, L., 2013. Effect of nitric oxide on antioxidative response and proline metabolism in banana during cold storage. J. Agric. Food Chem. https://doi.org/10.1021/jf401447y.
Xie, C., Chu, B., He, Y., 2017. Prediction of banana color and firmness using a novel wavelengths selection method of hyperspectral imaging. Food Chem. 245, 132–140. https://doi.org/10.1016/j.foodchem.2017.10.079.
Xie, C., Chu, B., He, Y., 2018. Prediction of banana color and firmness using a novel wavelengths selection method of hyperspectral imaging. Food Chem. https://doi.org/10.1016/j.foodchem.2017.10.079.
Zhuang, J., Hou, C., Tang, Y., He, Y., Guo, Q., Miao, A., Zhong, Z., Luo, S., 2019. Assessment of external properties for identifying banana fruit maturity stages using optical imaging techniques. Sensors. https://doi.org/10.3390/s19132910.