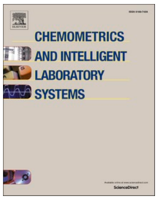
Chemometrics and Intelligent Laboratory Systems 263 (2025) 105446

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics



Comparison of colour and texture feature extraction methods to predict anthocyanins content in *Sangiovese* grapes

Camilla Menozzi [a], José Manuel Prats-Montalbán [b], Rosalba Calvini [a,c,*], Alessandro Ulrici [a,c]

[a] Department of Life Sciences, University of Modena and Reggio Emilia, Via Amendola, 2, Reggio Emilia, Italy
[b] Multivariate Statistical Engineering Group, Department of Applied Statistics, Operations Research and Quality, Universitat Politècnica de València, Camino Vera s/n, València, Spain
[c] Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali (INSTM), Via Giusti, 9, 50121, Firenze, Italy

ARTICLE INFO

Keywords:
RGB imaging
Colour
Texture
Data reduction
Data fusion

ABSTRACT

Colour and texture are the two main sources of information contained in RGB images of food products. Different image-level approaches are available to analyse the image properties based on the extraction of colour and texture features, and the selection of the most appropriate method is a critical point, since it could significantly impact the outcomes. The present study has three main objectives. Firstly, we propose an innovative data dimensionality reduction method to extract and codify the texture features of an RGB image into a one-dimensional signal, named texturegram (TXG). Then, TXG approach is compared with different image-level feature extraction methods, such as colourgrams (CLG), Soft Colour Texture Descriptors (SCTD) and Grey Level Co-occurrence Matrices (GLCM). These techniques were used to analyse a benchmark dataset of RGB images already considered in a previous study to build Partial Least Squares (PLS) models and relate the image features with anthocyanins content of red grape samples. We also investigated the possible advantages of combining the colour and texture information brought by the different image-level techniques using data fusion. PLS models were calculated considering different partitions of the RGB image dataset into training and test set. The performances of the different models were statistically evaluated by means of Analysis of Variance (ANOVA) and Principal Component Analysis (PCA). Overall, the results suggested an interesting, even if slight, improvement of the model performances when fusing CLG and TXG, but also highlighted the hybrid nature of TXG to simultaneously explore colour and texture properties.

1. Introduction

The use of Red, Green, Blue (RGB) digital imaging has rapidly increased in the field of analytical chemistry to inspect food quality [1–4]. As a matter of fact, the visual aspect of food plays a significant role in quality assessment and, at the same time, automated analyses are highly needed for food monitoring. The attractiveness of RGB imaging lies in its ability to provide an objective evaluation of food products through technically sophisticated systems that are accessible at relatively low cost.

Colour and texture are two of the most important sources of information contained in RGB images. They are inextricably linked and are always present in an image, even if one property can sometimes prevail over the other. In a greyscale image, the concept of tone is based on the varying shades of grey of the pixels in an image, while texture is concerned with the spatial distribution of grey tones [5]. The same principles can be extended to RGB images, where colour corresponds to pixels intensity values in the red, green and blue channels, while texture can be defined as the spatial variation of pixels colour intensities, i.e., the dependence between pixels and their neighbouring ones in terms of colour variation [4].

RGB image analysis involves processing colour-related and texture-related information of RGB images, and it can be performed at two different levels, i.e., at the pixel-level and at the image-level.

Pixel-level approaches consider each pixel of an image as a separate sample that can be characterized by its colour, intensity or location, and these properties can be simultaneously evaluated according to Multivariate Image Analysis (MIA) [6].

On the other hand, image-level approaches consist of extracting different types of colour or texture features from images, to provide an

---

* Corresponding author. Department of Life Sciences, University of Modena and Reggio Emilia, Via Amendola, 2, Reggio Emilia, Italy.
E-mail address: rosalba.calvini@unimore.it (R. Calvini).

https://doi.org/10.1016/j.chemolab.2025.105446
Received 7 February 2025; Received in revised form 26 May 2025; Accepted 26 May 2025

appropriate image characterisation and to consider the whole image as a single object. In this way, the information of interest contained in each image is encoded into a feature vector containing colour and/or texture descriptors. This approach is particularly useful when it is necessary to simultaneously compare a large number of images and to gain a deep understanding of the whole dataset structure. As a matter of fact, in practical applications, since each image is often supposed to enclose the properties of a given sample, the possibility to get an immediate and overall estimate of a sample is preferable, focusing on the single pixels only when necessary [7].

Digital image analysis can also be performed with Deep Learning (DL) approaches that learn feature representations directly from data, allowing to extract and model intricate image features in a blind and automated way. This property is due to the ability of Artificial Neural Networks (ANNs) architectures to inherently discern pertinent features, bypassing the need of explicit feature extraction of traditional MIA methods [8]. Even if the automated feature extraction of DL approaches can be seen as a significant advantage over more traditional MIA methods, it has to be considered that the robustness and reliability of DL models depend on the availability of copious data to train intricate ANNs network architectures, characterized by a strongly multilayered composition [8,9]. However, in practical food-related applications, it is often not feasible to acquire the adequate number of samples needed to train a DL model. Indeed, in the quantification of chemical compounds it is necessary not only to collect samples and acquire their images, but also to perform the reference chemical analyses in the laboratory. This latter step is often laborious, time consuming and expensive, therefore limiting the number of samples that can be considered for model training. Furthermore, feature extraction in DL methods is performed in a totally blind way by the algorithm, making it difficult the final interpretation of the model. Therefore, MIA and DL are two distinct approaches to image analysis, each offering different advantages and disadvantages, and catering to different fields of application. Even if MIA can be considered a more traditional approach compared to DL, both methods are widely used in current computer vision applications for the analysis of food matrices [9–13]. Due to the nature of the dataset considered in this study, we decided to focus the investigation on image-level MIA approaches to extract colour and texture features.

Colour features can generally be extracted from different colour spaces, of which RGB (red, green, blue) and HSI (hue, saturation, intensity) are the most common [14]. In this context, some of the authors have previously proposed the colourgrams method as an image-level approach accounting for colour information. Based on this method, each image is transformed into a *colourgram* (CLG), which is a one-dimensional signal of 4900 points encoding the colour information of the corresponding image using the frequency distribution curves of several colour-related descriptors, including R, G and B values, the relative colours, H, S and I values, and the score values of different Principal Component Analysis (PCA) models calculated on RGB data [15–19]. Hence, the colourgrams approach allows a comprehensive investigation of many colour-related properties of an image together with a significant reduction of image data dimensionality, enabling the possibility to handle large datasets of RGB images with a relatively low computational effort. This represents a great advantage of colourgrams approach, since several colour-related parameters are simultaneously evaluated without the need of defining *a priori* which are the most suitable for the problem at hand. It is then possible to select the most relevant colour descriptors *a posteriori* using variable selection methods. A simplified version of colourgrams has been recently proposed in literature based only on some colour histograms (i.e., R, G, B, L, H, S, I or a subset of these parameters) and excluding the part of the signal obtained through the data elaboration by PCA [20]. However, this simplified approach requires available information about the best possible combination of colour parameters for a given problem [20]. Furthermore, different applications of the colourgrams approach demonstrated the relevance and usefulness of the PCA part of the signal [21,22].

Concerning the extraction of image texture features, the available approaches can be grouped into four main categories: statistical methods, structural methods, transform-based methods and model-based methods [23].

Statistical methods are the most common ones due to the relative simplicity and immediacy associated with these techniques. Based on the assumption that the spatial distribution of grey levels in an image is one of the qualities defining texture, statistical methods involve the extraction of a set of statistical parameters from such distributions. Statistical methods can be further classified into first-order, second-order and higher-order statistics [24].

First-order statistics describe the distribution of individual pixel values, independently of the spatial relationship between pixels [25–27]. This category encompasses the calculation of Soft Colour-Texture Descriptors (SCTD), a set of statistical features describing colour and texture in a defined colour space. As an example, SCTD can include the mean and 2nd to 5th order mathematical central moments of the single R, G and B channels or of colour parameters of colour spaces different from RGB [28,29].

Second-order statistics are more suitable to describe non-random textures since they account for the spatial interaction between pairs of pixels in an image [6,30]. The most common second-order statistics are the texture features that can be extracted from Grey Level Co-occurrence Matrices (GLCM), such as contrast, correlation, energy, homogeneity and entropy among others [6,31–33].

Nevertheless, the search and the development of other approaches for texture evaluation has not yet come to an end. In this work, we propose an alternative statistical method to extract texture features from RGB images, which consists of codifying the image texture properties into a one-dimensional signal, hereinafter named *texturegram* (TXG). The proposed approach is based on a combination between a modified version of the approach proposed by Ref. [34] and lately extended to RGB images by Ref. [35], and a data dimensionality reduction algorithm of hyperspectral images proposed by Ref. [7].

According to Bharati and MacGregor [34] each pixel of a grey level image can be characterized not only by its intensity values but also by the intensity values of neighbouring pixels. In this manner it is possible to capture the spatial information associated with each pixel by linking its grey level value with the grey level values of neighbouring pixels. Prats-Montalban and Ferrer [35] applied the same approach to the three R, G and B channels composing an RGB image in order to obtain textural information from such data.

In this study we applied a modified version of the above-mentioned methods, since for each channel we account for texture information by calculating for each pixel the differences between its intensity values and the intensity values of neighbouring pixels in some specific direction. The resulting outcome can be seen as a multichannel image where each channel corresponds to the pixel intensity differences calculated for a given spatial direction in one of the original RGB channels. In order to retrieve useful information from such multichannel images we applied the Common Space Hyperspectrograms (CSH) approach, an algorithm initially proposed to perform data dimensionality reduction of hyperspectral images [7,36]. The principle behind CSH consists of converting each hyperspectral image into a one-dimensional signal, the hyperspectrogram, obtained by merging in sequence the frequency distribution curves of quantities (i.e., scores, Hotelling T^2 and Q residual values) obtained from a PCA model common to all the images of the dataset. Since the multichannel images obtained by computing the differences between neighbouring pixels in the original RGB domain account for texture information, in this case the signals obtained from CSH approach are referred to as *texturegrams*. Therefore, each texturegram acts like a fingerprint of texture properties of the corresponding RGB image. To our knowledge, this is the first time that the method originally proposed by Bahrati and MacGregor is coupled with CSH approach to describe texture properties of images at the image-level.

When evaluating food quality, it is difficult to establish the most appropriate method for image feature extraction given the very wide variety of food products and issues to deal with. The selection of the most relevant features to be analysed is a critical point, since it could have a significant influence on the outcomes. Different techniques allow focusing on different features and one or several methods might be employed together to acquire image features for food quality evaluation [4].

In this work we applied different image-level techniques to analyse colour and texture features of a benchmark dataset of RGB images already considered in a previous study [37], where the colourgrams approach has been successfully applied to predict anthocyanins content of red grape samples from RGB images of grape berries acquired with a smartphone-based device. Although in this type of dataset colour is the primary aspect to consider when assessing anthocyanins content, we believed that texture-related information could also help to improve the predictive ability of the calibration models. For example, the presence of the green pedicels of the grape berries, which have the same colour as unripe berries but different shapes, and/or the image background, could negatively affect the model performance if only colour is considered.

The aim of this work is to compare the results obtained from different feature extraction methodologies, namely based on colourgrams (CLG), Soft Colour-Texture Descriptors (SCTD), second order statistics extracted from co-occurrence matrices (GLCM) and texturegrams (TXG).

While CLG are a colour-based feature extraction method, SCTD, GLCM and TXG enable to code in a different manner the texture properties of images. TXG approach is firstly presented in this study, and its performances are compared with SCTD and GLCM, two well-established image-level texture feature extraction methods. Both SCTD and GLCM are based on coding texture information using statistical descriptors calculated from the images, but they account for different texture aspects. Indeed, SCTD are effective in describing well-spread random textures, while GLCM is effective in describing specific and regular patterns [28,29,38].

Furthermore, we also investigated the possible advantages and usefulness of combining the colour and texture information of the dataset by low-level and mid-level data fusion.

2. Materials and methods

2.1. Samples

In this work, we considered grape samples of *Sangiovese* grape variety, as it is the most commonly cultivated red grape variety used for wine-production in Italy [39].

Samples were collected in the province of Ravenna (Italy). Grape samples were collected at different ripening times, from about 30 % veraison up to complete ripeness during vintages 2020 and 2021. Two different vineyards were involved in vintage 2020 (Vin A and Vin B), while an additional vineyard (Vin C) was included in 2021 vintage for a total of 3 vineyards. The three vineyards were located in different areas and 10 harvest times were considered for each of them in each vintage, for a total of 50 grape samples (i.e., 20 samples in vintage 2020 + 30 samples in vintage 2021).

Sampling of the grape berries was performed to be as representative as possible of the main sources of variability that should be taken into account when monitoring grape ripening, i.e., different vineyards, different years, and different plants at different ripening times.

For each grape sample six field aliquots were collected and separately imaged (*see* Section 2.3).

2.2. Measurement of total anthocyanins

Total anthocyanins content (TAnt) was determined according to Glories method [37,40,41]. For each sample, grape berries were put in a blender (Kenwood CH-58) and mashed to homogenize pulp and skin. One half of the so-obtained grape mash was macerated in a pH 1 solution to have complete extraction of anthocyanins. After 4 h of maceration at room temperature under low agitation, the solution was filtered on glass wool. Then, a spectrophotometric determination of anthocyanins (JASCO V-530, Tokyo, Japan) was performed on the clear juice using the methodology based on bleaching with sulphur dioxide [42–44] after which the total anthocyanins content (TAnt, expressed as mg/kg of malvidin-3-glucoside) was determined.

2.3. Image acquisition

The RGB images of the grape samples were acquired with a portable smartphone-based device specifically developed for on-field determination of grape phenolic maturity [37]. Briefly, the acquisition device consists of a black case made of acrylonitrile butadiene styrene (ABS) with size 16 cm × 10 cm × 15.5 cm. The grape berries were placed inside a drawer (sample holder) positioned in the lower part of the acquisition device together with a cardboard with colour references, while the smartphone was positioned in a specific case located in the sliding lid on the top of the device. Controlled illumination conditions were ensured by a strip of white light emitting diodes (LED) placed inside the case and powered by a power bank fixed outside the device.

Image acquisition was performed using a dedicated app installed on a Xiaomi Redmi Note 8 smartphone (Xiaomi, China), equipped with a 48 megapixels camera with CCD technology, a spatial resolution of 8000 × 6000 pixels, and considering a shutter speed of 1/50 s, f/1.9 lens aperture and ISO 200. A cardboard of standard colour references was included in each image scene, to perform image correction (*see* Section 2.4).

Each one of the 6 aliquots of a sample was separately imaged, for a total of 300 acquired images (=6 aliquots × 50 grape samples). The previous analysis of this dataset [37] allowed identifying as outlier one grape sample that was removed. Therefore, in this study we excluded from further elaborations the 6 aliquot images obtained from the outlier sample and only the remainder 294 images were kept (=49 samples × 6 aliquots). The final dataset of 294 images corresponded to a file size of 85 MB.

For more details about the image acquisition system and the smartphone app, the reader is referred to Ref. [37].

2.4. Image correction

As described in the previous section, the scene of each image was composed by two different areas: the area with the cardboard of Standard Colour References (SCR) and the sample area with the grape berries, as shown in Fig. S1 of Supplementary Material. The SCR consisted of 12 rectangular patches of different colours printed with a professional printer on a non-reflective photographic paper. After being properly selected and cropped from the original image, the SCR area was used to perform an internal image calibration. This procedure is highly recommended to minimize the possible variability among images, principally owed to instrumental instability over time.

The first step of image correction consisted of the selection of an image referred to as *master image*, which is an image of the dataset whose SCR area was considered as reference for the further correction steps. In fact, the SCR of the master image was used to correct all the other images, referred to as *slave images*.

The correction algorithm was based on an internal calibration performed channel by channel, between median vectors calculated for red (R), green (G) and blue (B) channels of the SCR of the *master* image and the median vectors of the SCR of each remainder *slave* image. The SCR of each slave image was corrected to be as similar as possible to the SCR of the master image only when significant differences were detected; in this case, the same correction function was applied to correct each pixel of the corresponding sample area. After some preliminary tests, a first-order polynomial was considered for image correction.

The standardization procedure was implemented in the software of the device used for image acquisition, to automatically correct the acquired images when necessary.

For a more detailed description about the algorithm employed to standardize the images, the reader is referred to Ref. [22]; the image correction algorithm is available in RGB Correction GUI, a graphical user-friendly interface working under MATLAB environment (The MathWorks Inc., USA) and downloadable for free from https://www.chimslab.unimore.it/downloads/.

3. Image-level extraction of colour and texture features

After image correction, the sample area (i.e., the area containing the grape berries) was cropped from the corrected images obtaining smaller images of size 2310 × 2623 pixels. Only the cropped images containing the grape berries were considered for further elaboration, which consists of the extraction of colour and texture features using different techniques, i.e., colourgrams (CLG, *see* Section 3.1), texturegrams (TXG, *see* Section 3.2), Soft Colour Texture Descriptors (SCTD, *see* Section 3.3) and Grey Level Co-occurrence Matrices (GLCM, *see* Section 2.5.4). All the considered approaches perform an image-level extraction of the features of interest, which consists of calculating for each image a feature vector accounting for colour or texture descriptors. These feature vectors obtained from all the images of the dataset can be collected into a data matrix, which in turn can be analysed using unsupervised or supervised chemometric methods.

3.1. *Colourgrams*

The cropped images of the 6 aliquots of the same sample were concatenated into a single image with size 2310 × 15738 pixels (hereinafter named *scanset*, *see* Fig. S2 of Supplementary Material) to better account for sample variability within the same ripening stage and thus obtain more robust and reliable results when developing the calibration models. Overall, a dataset including 49 *scanset* images was obtained.

Each *scanset* can be converted into the corresponding colourgram, i.e. a one-dimensional 4900 points-long signal, codifying the colour-related information of an RGB image and allowing significant data compression.

At first, each RGB image is unfolded in a two-dimensional matrix which contains all the pixels in the rows and the R, G, B channels in the columns [21]. Then, 16 additional colour-related parameters are calculated for each pixel and derived from the R, G and B values: lightness ($L = R + G + B$), relative colours ($rR = R/L$; $rG = G/L$; $rB = B/L$), the hue, saturation and intensity values of the corresponding HSI colour space, and the 3 PCA score vectors calculated on the RGB matrix considering raw ($SC1_R$, $SC2_R$, $SC3_R$), mean centred ($SC1_M$, $SC2_M$, $SC3_M$) and autoscaled data ($SC1_A$, $SC2_A$, $SC3_A$), as there is no *a priori* knowledge of which pretreatment performs best. For each PCA model, loading vectors and eigenvalues are calculated too. Therefore, 19 colour-related parameters are considered: the 3 R, G and B channels and the 16 additional parameters mentioned above.

For each colour-related parameter, the corresponding frequency distribution curve (i.e., histogram) is calculated considering 256 bins. Notably, even if the L, rR, rG and rB parameters are simple combinations of the R, G and B channels, they reflect different colour-related aspects, thus their histograms are not necessarily correlated to the R, G and B ones. Then, after normalising by the number of pixels, the frequency distribution curves of each *scanset* image are merged in sequence to obtain a one-dimensional signal which composes the first part of the colourgram. Finally, the loading vectors and eigenvalues of the 3 PCs for the 3 PCA models calculated on RGB data are added at the end of the colourgram to complete the signal.

Following the procedure described above, each *scanset* image was converted into the corresponding colourgram and the signals were collected into a data matrix with size 49 × 4900, where each row corresponds to the colourgram derived from a specific *scanset* image. The CLG dataset was stored in a file with size equal to 1 MB.

The colourgrams matrix can be analysed by multivariate analysis techniques enabling to evaluate all the images together and consider the colour-related information of the dataset as a whole. Even if the computation of colourgrams implies to calculate multiple parameters from R, G and B values, the procedure is easily automatable and the computational time required is quite low, allowing an effective and comprehensive characterisation of food samples without requiring *a priori* knowledge of the properties of interest for the given task.

Fig. 1 shows the colourgrams matrix obtained in this study, where each signal is coloured according to the anthocyanins content of the corresponding grape samples, and the vertical dashed lines define the colourgram regions related to the histograms of the 19 colour-related parameters, identified with the corresponding abbreviations.

For a deeper insight of the algorithm developed to create the colourgrams, the reader is referred to Refs. [15,19].

Colourgrams were calculated using Colourgrams GUI (Calvini et al., 2020), a graphical user interface written in MATALAB language (The Mathworks Inc., USA) and freely downloadable from https://www.chimslab.unimore.it/downloads/.

3.2. *Texturegrams*

In this work, we propose an innovative method to codify the texture features of an image, which is based on the same idea of colourgrams: converting each image into a one-dimensional signal, i.e., the *texturegram*, retaining the relevant information contained in the corresponding original image. As suggested by their name, the texturegrams take into account the texture-related properties of RGB images. The procedure used for the calculation of texturegrams involves two main steps: the conversion of each RGB image into a multichannel texture image accounting for texture properties and the application of a data dimensionality reduction algorithm to convert the multichannel texture images into one-dimensional signals.

3.2.1. *Conversion of RGB images into multichannel texture images*

In the first step of texturegram calculation, each original RGB image is converted into a multichannel image accounting for texture information using a modified version of the approach initially proposed by Ref. [34] and lately extended to RGB images by Ref. [35].

More in detail, for each channel of a single RGB image, the spatial-textural information associated with each pixel is obtained by linking its intensity value with those of neighbouring pixels at a given distance *d* [34]. Considering for example the red channel, the corresponding grey scale image can be unfolded into a column vector with as many elements as the number of pixels and each element corresponds to the red value of a single pixel. For each pixel it is possible to register also the red values of neighbouring pixels at distance *d* selected according to an established order, i.e., the pixel above, then in the right-up corner, then on the right and so on, following the clockwise direction. In this manner, each pixel is characterized by 9 values: its red value and the red values of the 8 neighbouring pixels. This information can be calculated across all the image pixels obtaining a textural matrix with 9 columns, corresponding to the intensity value of each pixel and the intensity values of neighbouring pixels in the 8 considered directions.

The same procedure can also be extended to the remainder G and B channels of an RGB image [35] and the textural matrices obtained for each colour channel can be subsequently stacked one beside the other, obtaining a colour-textural RGB matrix with 27 columns (=9 columns for each channel × 3 channels).

In this study, to extract the texture features from each RGB image we used the same principle described above but in a modified version. In fact, after unfolding, for each channel instead of reporting the values of the considered pixel together with those of the neighbouring pixels, we described the texture properties of each pixel by calculating the

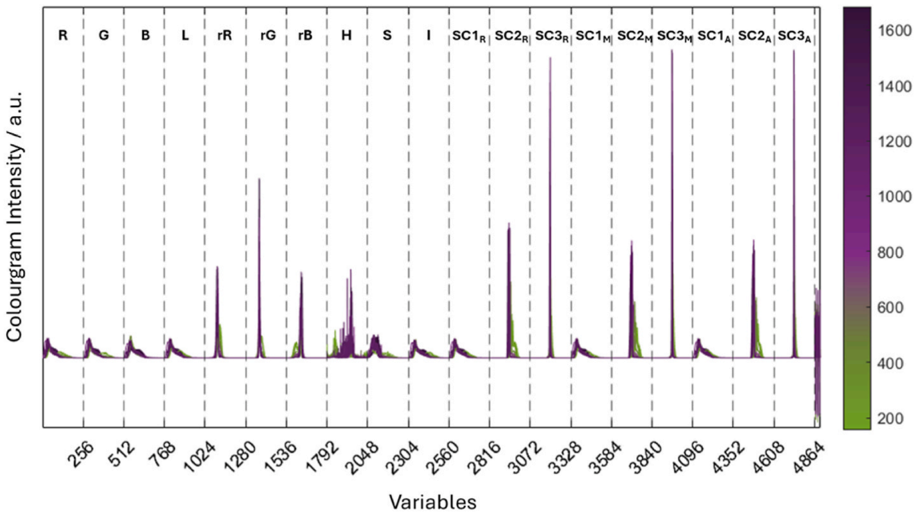


Fig. 1. Colourgrams calculated starting from the scansets; each signal is coloured according to the total anthocyanins content of the corresponding sample.

difference between its intensity value and the intensity values of neighbouring pixels in all the 8 possible directions, following a clockwise order starting from the pixel above the considered pixel. When calculating the differences between the reference pixel and its neighbours in the different orientations, the sign was kept. As a result, for each channel we obtained a two-dimensional matrix with 8 columns, corresponding to the 8 spatial directions, and a number of rows which is lower than the number of pixels of the original image and depends on the considered distance between neighbouring pixels. Indeed, for a given distance *d*, we decided to discard the information related to nominal pixels located at a distance lower than *d* from the edges of the original image since they lack some neighbours. The number of rows (*r)* of this bidimensional

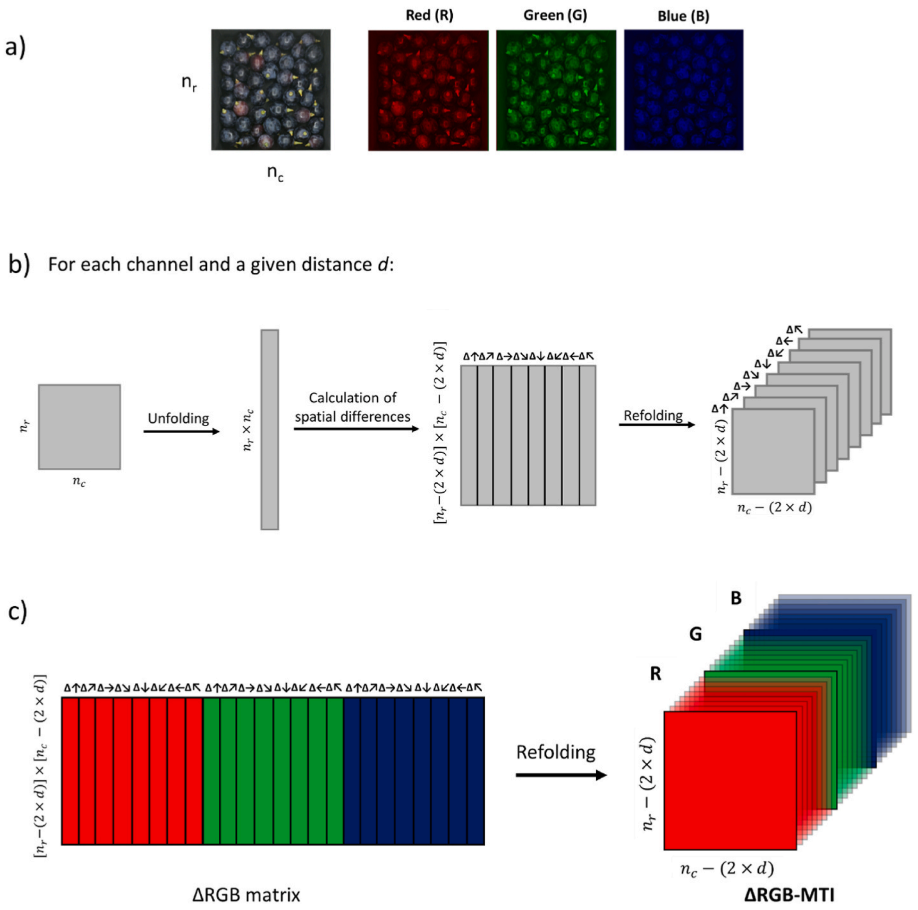


Fig. 2. Creation of a ΔRGB-MTI from an image of an aliquot: a) extraction of the red (R) green (G) and blue (B) channels; b) for each colour channel, calculation of the corresponding multichannel texture image accounting for the intensity differences between nominal pixels and the corresponding neighbours in the 8 different directions at a given distance d and c) ΔRGB-MTI obtained from the concatenation of the matrices obtained in b). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

matrix can be calculated as follows:

$$r = [n_r - (2 \times d)] \times [n_c - (2 \times d)] \quad (1)$$

where n_r and n_c represent the number of row and column pixels of the original image, respectively.

The bidimensional matrix accounting for pixel differences can be refolded back into the original image domain obtaining a multichannel texture image with $n_r - (2 \times d)$ rows, $n_c - (2 \times d)$ columns, and 8 channels (Fig. 2b).

This procedure was applied separately to all the three R, G and B channels of an image and the corresponding bidimensional difference matrices were staked one beside the other obtaining a new matrix, from here onwards named as ΔRGB matrix. Each ΔRGB matrix is a two-dimensional matrix with *r* rows (*see* Eq. (1)) and 24 columns, corresponding to 8 directions × 3 channels.

Then, the ΔRGB matrix is refolded back into the original image domain obtaining a ΔRGB Multichannel Texture Image (ΔRGB-MTI), with $n_r - (2 \times d)$ rows, $n_c - (2 \times d)$ columns and 24 channels (Fig. 2c). The procedure used to obtain the ΔRGB-MTI from RGB images is summarized in Fig. 2.

In this manner, each RGB image is converted into the corresponding ΔRGB-MTI, a multichannel image accounting for texture properties by considering the colour differences between each pixel and its neighbours. These ΔRGB-MTIs can be analysed using the common MIA methods available for pixel-level exploratory analysis of multispectral and hyperspectral images.

However, when dealing which a large number of images altogether that need to be analysed simultaneously, it is necessary to work at the image level and apply proper data dimensionality reduction algorithms, which is the second step in texturegrams calculation.

3.2.2. Conversion of ΔRGB multichannel texture images into signals

In the second step, the ΔRGB-MTIs are converted into the corresponding texturegrams using the Common Space Hyperspectrograms (CSH) algorithm, a data dimensionality reduction method originally proposed for the simultaneous analysis of hyperspectral images [7]. The main idea behind the CSH approach is to convert each hyperspectral image into a hyperspectrogram, a one-dimensional signal accounting for the relevant information obtained by merging in sequence the frequency distributions of the score vectors, Q residuals and Hotelling T^2 values obtained from a PCA space common to all the images of the dataset. For a detailed description of the CSH approach and the algorithm used to compute the common PC space the reader is referred to Ref. [7].

The different steps in the computation of texturegrams using the CSH approach can be summarized as follows:

1. Pixel-wise mean centring of the ΔRGB-MTIs according to the grand mean, i.e. to the mean calculated for each variable (thus, for each of the 8 directional differences of intensity in each of the three colour bands) on all the retained pixels.
2. Calculation of the kernel variance-covariance matrix (Z) of the whole dataset. Z has size {24 × 24}, where 24 is the number of channels of the ΔRGB-MTIs. To make this step computationally feasible, the variance-covariance matrices were separately calculated for each ΔRGB-MTI and subsequently summed entry-wise to obtain Z.
3. Decomposition of Z by singular value decomposition (SVD) to obtain the loading vectors of the common PC space. After some preliminary tests, in this work we decided to keep 16 PCs.
4. Calculation of the scores, Q residuals and Hotelling T^2 vectors of each ΔRGB-MTI by projecting the ΔRGB-MTIs onto the common PC space.
5. For each ΔRGB-MTI, computation of the frequency distribution curves for each score, Q residuals and Hotelling T^2 vector. For each vector, the corresponding frequency distribution curve was calculated considering a common range corresponding to the interval between the global minimum and the global maximum of the outputs of the PCA model obtained for all the images. After some preliminary tests, we considered 100 bins to obtain the frequency distribution curves (i.e., histograms) and each histogram was normalized according to the number of pixels of the corresponding image.
6. Creation of the texturegrams by joining in sequence the frequency distribution curves calculated in step 5.

In this work, each texturegram resulted in a 1800 points-long signal, resulting from 18 frequency distribution curves (16 PCs + Q residuals + Hotelling T^2) × 100 bins.

The procedure used to convert the ΔRGB-MTIs into texturegrams is schematically represented in Fig. 3 where, for the sake of conciseness, only 2 PCs are considered to calculate the common PC space.

According to the approach proposed in this paper, a whole dataset of RGB images can be converted into a matrix of texturegrams, in which each row represents a sort of texture-fingerprint of the imaged sample. The texturegrams matrix can be easily processed by means of different multivariate statistical methods for exploratory analysis or for the calculation of supervised regression or classification models. In other words, each signal codifies a significant part of the texture features of interest contained in a specific image and it is possible to make quick comparisons of a high number of images at the same time.

As will be better detailed later in Section 4, in this study the PLS models were obtained after splitting the RGB image dataset into training and test images using different partitions. To ensure a real external validation of the models, the computation of the grand mean (step 1) and of the common PC space (step 2 and step 3) were performed considering only the ΔRGB-MTIs obtained from the training set images. When dealing with the test set images, the corresponding ΔRGB-MTIs were firstly mean centred according to grand mean of training data and then projected into the PCs space common to the whole training dataset (step 4). Then the texturegrams of the test set were calculated following steps 5–6.

In this study, the texturegrams were firstly calculated separately for each one of the 6 aliquot images of a *scanset*. Then, the 6 texturegrams of the aliquot images of the same sample were averaged in order to obtain a single texturegram for each *scanset*.

Furthermore, in the computation of texturegrams it is necessary to define the distance *d* of neighbouring pixels used to convert the original RGB images into the corresponding ΔRGB-MTIs. In this study we considered three different distance values, corresponding to 5, 10 and 15 pixels respectively, and for each distance a texturegrams matrix with size {49 × 1800} was obtained. Each TXG dataset corresponded to a file size of 0.44 MB. In addition to the distance between neighbouring pixels, TXG computation also requires defining the number of PCs and the number of bins of the frequency distribution curves composing the signals. In this study we decided to use fixed values for these parameters (i.e., 16 PCs and 100 bins), chosen based on preliminary tests. Conversely, varying distance values were considered to allow a direct comparison with GLCM (*see* Section 3.4).

As an example, Fig. 4 shows the texturegrams matrix obtained in this study considering a distance of 15 pixels, where each signal is coloured according to the anthocyanins content of the corresponding grape sample.

The conversion of RGB images into ΔRGB-MTIs was performed using a script written ad hoc in MATLAB language, while the conversion of ΔRGB-MTIs into signals was performed using Hyperspectrograms GUI, a graphical user interface working in MATLAB environment and available for the conversion of multispectral or hyperspectral images into CSH. Hyperspectrograms GUI is freely downloadable from https://www.chimslab.unimore.it/downloads/.

3.3. Soft Colour-Texture Descriptors

Soft Colour-Texture Descriptors (SCTD) [28,29] are a set of statistics requiring a relatively low computational demand to be extracted and

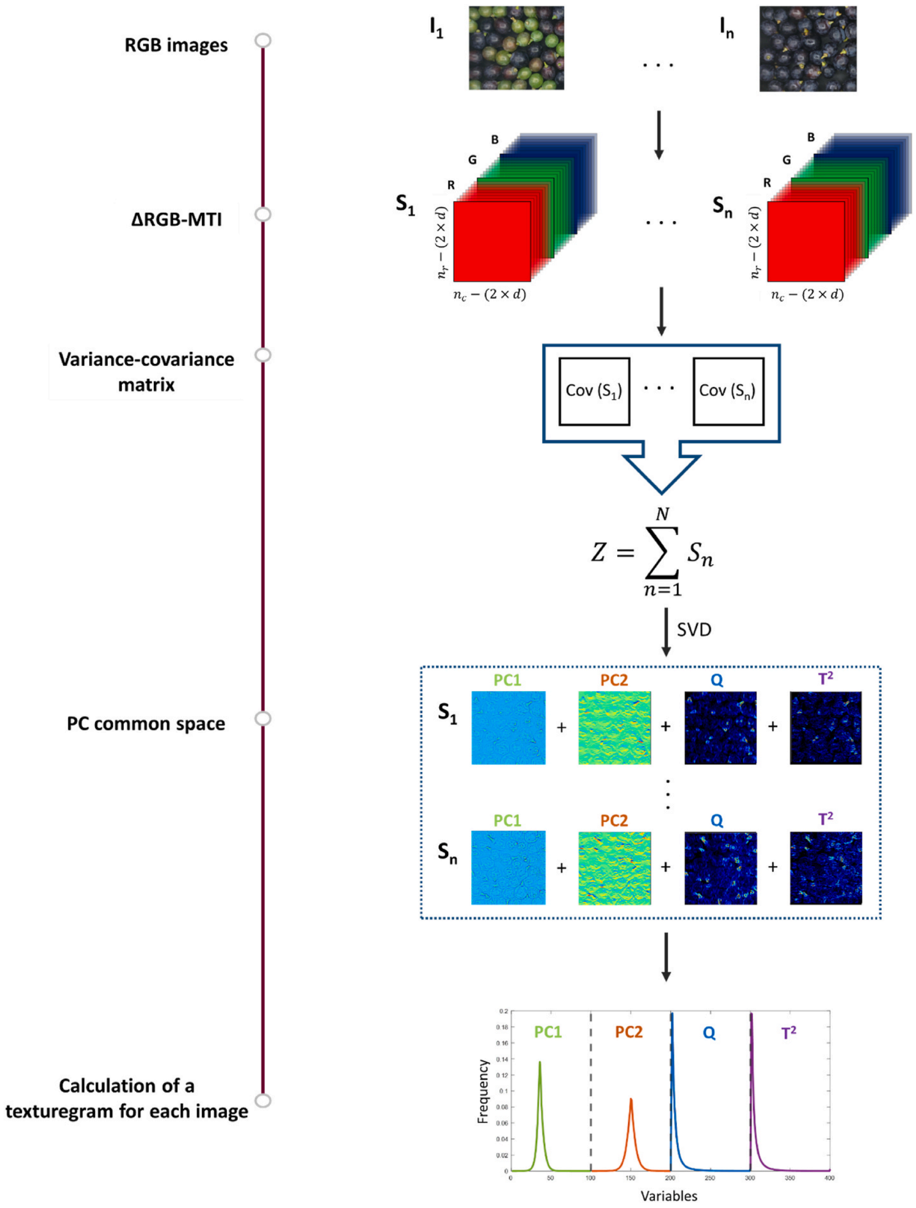


Fig. 3. Schematic overview of the whole process to calculate the corresponding texturegrams from RGB images. To simplify the visualization, in this scheme the common PC space includes only 2 PCs, while the TXG dataset considered in this study was obtained considering 16 PCs.

elaborated when approaching colour and texture description in RGB images.

In fact, these features can be extracted by computing the mean and the different mathematical statistical central moments from 2nd to 5th (i.e., variance, skewness, kurtosis and 5th order) related to each vectorized image channel, for a total of 15 descriptors in an RGB image, i.e., (1 + 4) statistical features × 3 colour channels.

In this work, the SCTD were firstly calculated separately for each aliquot image. Then, for the 6 aliquot images of a *scanset* the mean and variance of each descriptor were calculated to consider also the internal variability of each descriptor within the same sample. In this manner, for each *scanset* a feature vector of 30 parameters was obtained (15 mean values of SCTD + 15 variances of SCTD). The so-obtained SCTD matrix has dimensions of {49 × 30}, corresponding to a file size of 0.015 MB.

3.4. Grey level Co-occurrence matrices (GLCM)

A Grey Level Co-occurrence Matrix (GLCM) [5] is a matrix containing texture features of each image and it is obtained using second order statistics. For a grey-scale image, the GLCM has a number of both rows and columns equal to the discrete grey level intensities of the image in order to embrace all the possible combinations of intensities that can appear between two image pixels at a given relative position. The relative pixel position has to be defined *a priori* before calculating GLCM, usually by setting a certain distance (d) and a certain angle (θ).

A GLCM can be also seen as a cross tabulation of a bidimensional histogram associated to the random distribution of the bidimensional variable [45]. In fact, each element (i, j) of the GLCM expresses the total number of occurrences (i.e., the relative frequencies) for which an image pixel with intensity value i is coupled to another image pixel with intensity value j, at the given spatial relationship depending on the considered distance and angle. Several features can be extracted from the GLCM, although only a few are widely used and, in this study, we focused on five of the most common ones, namely contrast, correlation, energy (also known as Angular Second Moment), homogeneity and entropy [46].

When defining the spatial relationship among pixels to be considered, the distance of 1 pixel is the most commonly used, while four

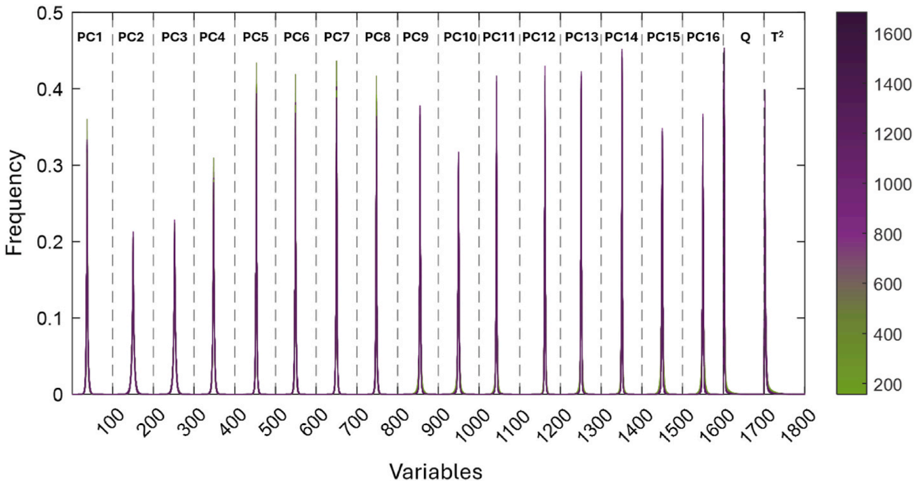


Fig. 4. Example of texturegrams calculated starting from the RGB images considered in this study and considering a distance of 15 pixels; each signal is coloured according to the total anthocyanins content of the corresponding sample.

directions are typically explored and correspond to angles of 0°, 45°, 90° and 135°, corresponding to four distinct GLCMs for each original grey-scale image. Usually, when the features sought are expected to be invariant to the direction, it is a common practice to average the statistical parameters extracted from the four GLCMs corresponding to the four angles.

All these principles can be easily extended to RGB images, by calculating GLCM features for each single channel.

In this work, GLCMs were firstly calculated separately for each aliquot image considering for each channel 256 intensity levels and the four directions listed above. For each channel, the values of the 5 statistical features (i.e., contrast, correlation, energy, homogeneity and entropy) extracted from the four GLCMs corresponding to the four directions were computed. Hence, for each aliquot image we obtained a final vector of 15 features (5 GLCM features × 3 colour channels).

Similarly to SCTD, for the 6 aliquot images of a single *scanset* the mean and variance of every GLCM parameter were calculated to consider the internal variability of every parameter within the same sample. In this manner, a final row feature vector of 30 parameters (15 average GLCM features + 15 variances of GLCM features) was obtained for each *scanset*.

This procedure was performed considering four distance values in GLCM computation corresponding to 1, 5, 10 and 15 pixels, which led to four separate datasets of GLCM descriptors with size {49 × 30}. Each GLCM dataset was stored in file with size equal to 0.015 MB.

4. PLS calibration models

PLS calibration models were calculated in order to relate the colour and/or texture features extracted from the RGB images of grapes to the anthocyanin content of the corresponding samples. More in detail, the PLS models were calculated not only considering the different image-level feature extraction methods as separate data blocks (*see* Section 4.1) but also combining the colourgram approach with each one of the texture extraction methods (*see* Section 4.2).

For the computation of the PLS models, multiple partitions of the acquired RGB images into training and test images were considered. In fact, in order to make a detailed and accurate evaluation of strengths and weaknesses of the feature extraction methods considered in this study, the PLS calibration models were calculated by iteratively including in the test set only the samples collected in a specific vineyard and in a specific year. For example, the partition “Vin_A_2020” in Table 1 refers to the training/test set partition in which the *scanset* images of grape samples collected in vineyard A in 2020 were used as test images, while the remainder images were used as training images. In this manner, 5 different training/test set partitions were considered. More detailed information about the considered training/test set partitions is reported in Table 1. It is interesting to highlight that in the partition “Vin_C_2021” the external validation considers as test set grape samples harvested in a vineyard not considered during calibration. We decided to consider these multiple TR/TS partitions that iteratively use as TS set the samples collected in a specific vineyard and/or year in order to assess the prediction performances against different scenarios that may occur in practical applications, such as the need of performing predictions based on models trained on samples of previous years or collected in different geographical areas.

Table 1
Information about the different training/test set partitions considered in this work. The partitioning names stand for the samples collected in a specific vineyard and year iteratively left in the external test set, while “Original” refers to the training/test partition considered in the previous work.

TR/TS partition	N° training set samples	N° test set samples
Vin_A_2020	39	10
Vin_B_2020	39	10
Vin_A_2021	39	10
Vin_B_2021	39	10
Vin_C_2021	40	9
Original	34	15

In addition, the same training/test set partition approach adopted in the previous work [37] was also considered for the sake of comparison. This training/test set partition is referred to as “Original” in Table 1 and from here onwards in the manuscript. It consists of attributing to the test set the *scansets* of different harvest times for each vineyard as follows: 34 *scansets* in the training set and 15 *scansets* in the test set. Originally, this training/test set partition was carried out with the purpose of encompassing in the model the possible different trends expected due to the sampling procedure: even within the same grape variety, the different geographical position of the vineyards and the climatic variability from one vintage to another could significantly impact the accumulation of anthocyanins in grape berries during ripening.

The proper number of Latent Variables (LVs) was selected by minimizing the Root Mean Square Error in Cross-Validation (RMSECV).

According to the different training/test set partitions, specific cross-validation approaches were applied. When considering the “Original” training/test set partition, venetian blinds cross-validation with 5 deletions groups was performed. For the remainder partitions, a “one-vineyard-one-year-out” cross-validation was followed. In other words, for each training set, cross-validation was performed by considering 4 deletion groups, each one corresponding to the samples collected in only one vineyard in a specific year.

The results of the calibration models were expressed in terms of Root Mean Square Error calculated in cross-validation (RMSECV) and prediction of the external test set (RMSEP). Finally, the evaluation of the performance of the calibration models obtained using the different feature extraction techniques and their combinations was done by using both Analysis of Variance (ANOVA) and Principal Component Analysis (PCA).

All calculations were performed using the PLS Toolbox ver. 8.8.1 (Eigenvector Research Inc.,USA), the Statistics and Machine Learning Toolbox ver. 12.2 (The MathWorks Inc., USA), and ad hoc written functions running in the MATLAB ver. 9.11 environment (The Mathworks Inc., USA).

4.1. Single blocks

Initially, the PLS calibration models were calculated on each individual dataset accounting for colour or texture features of the original RGB images, i.e., colourgrams (CLG), texturegrams (TXG), Soft Colour Texture Descriptors (SCTD) and GLCM features (GLCM).

The dataset of colourgrams was pre-processed using mean centering, as in the previous study. As for texturegrams, three different datasets were obtained based on the distance value used for neighbouring pixels and corresponding to 5 pixels (TXG D5), 10 pixels (TXG D10) and 15 pixels (TXG D15). The PLS models were calculated for each TXG dataset using Pareto scaling and mean centring as preprocessing methods. Only one dataset of SCTD was obtained and it was pre-processed using autoscaling. Similarly to TXG, also for GLCM different distance values were considered, resulting in four GLCM datasets corresponding to distances equal to 1 pixel (GLCM D1), 5 pixels (GLCM D5), 10 pixels (GLCM D10) and 15 pixels (GLCM D15). The four GLCM datasets were pre-processed using autoscaling before the calculation of PLS models.

On the whole, 54 PLS models were calculated considering the single blocks, corresponding to 6 training/test set partitions × 9 single data blocks (1 CLG + 3 TXG + 1 SCTD + 4 GLCM).

4.2. Data fusion

In order to evaluate the possible advantages of integrating together the colour and texture information obtained from the different approaches starting from the same images, data fusion strategies were used to iteratively combine the information brought by CLG with each one of the texture analysis approaches described in the previous sections. Both low-level data fusion, which consists of merging together the different blocks of data, and mid-level data fusion, which consist of combining features extracted from the original data blocks, were considered [47].

4.2.1. Low-level data fusion

Considering low-level data fusion, the single blocks were separately pre-processed as described in Section 4.1 and then concatenated into a unique data matrix, with a number of rows equal to the number of samples, and a number of columns equal to the sum of the variables from each block. Consequently, the size of the low-level fused matrices was equal to 49 rows (samples) and:

- 6700 columns when fusing CLG and TXG data blocks (=4900 CLG variables + 1800 TXG variables); three CLG + TXG fused matrices were obtained, corresponding to the three distance values used to calculate TXG
- 4930 columns when fusing CLG with SCTD data block (=4900 CLG variables + 30 SCTD variables)
- 4930 columns when fusing CLG with GLCM data blocks (=4900 CLG variables + 30 GLCM variables); four CLG + GLCM fused matrices were obtained, corresponding to the four distance values used to calculate GLCM.

For each combination of colour- and texture-related data blocks, after low-level data fusion the resulting data matrix was pre-processed using block scaling before the calculation of the corresponding PLS model.

A total of 48 low-level data fusion models were calculated, corresponding to 6 training/test set partitions × 8 combinations between CLG and each one of the remainder 8 datasets accounting for texture features.

4.2.2. Mid-level data fusion

Mid-level data fusion consists of concatenating features of interest extracted from the different data blocks by a preliminary data analysis step. To this aim, the outcomes of the calibration models previously calculated on the single blocks (*see* Section 4.1) have been combined to perform mid-level data fusion, by concatenating the PLS score vectors. In particular, the score vectors obtained from the PLS model calculated on CLG were combined with the score vectors of each PLS model calculated on a data block of texture features. Therefore, the size of each mid-level data fusion matrix was equal to 49 rows (samples) and as many columns as the sum of the number of LVs of the PLS model calculated on the CLG block and the number of LVs of the PLS model of the other texture-related data block.

Considering mid-level data fusion, two different preprocessing strategies were tested on the fused datasets of scores vectors: autoscaling and mean centring followed by block scaling. When performing autoscaling all the single block PLS score vectors used as descriptors have the same scale, while with block scaling each block of data is scaled to unit variance. In other words, each block of data brings the same contribution to the calibration models regardless of the number of latent variables. At the same time, the relative scales of the score vectors within each block are kept [48].

On the whole, 96 mid-level data fusion models were calculated, corresponding to 6 training/test set partitions × 8 combinations between CLG and texture data × 2 preprocessing methods.

4.3. Statistical analysis of calibration performances

ANOVA was performed on the RMSEP values of all externally validated models with the samples collected in a specific vineyard and in a specific year, *i.e.*, not considering the models calculated using the “Original” training/test set partition (*see* Section 4). This was aimed at better focusing on the significance of the possible variability in the results due to the different vineyards in each vintage.

In our case, an ANOVA without interaction was performed on the RMSEP values obtained by the PLS models calculated on the single blocks, in order to highlight any statistically significant difference between the data coming from the single feature extraction techniques and the vineyards in each vintage. Regarding the models obtained using data fusion approaches, a new ANOVA with interaction was used to evaluate whether the effects of the techniques, the data fusion approaches, the vineyards and/or their interactions were significant.

Moreover, to obtain an overall comparison of the outcomes, PCA was carried out to compare the results coming from the different techniques and data fusion approaches, taking into account all the training/test set partitions applied in this study. PCA was calculated on the whole set of RMSECV and RMSEP values obtained considering both the single feature extraction techniques and their combinations through data fusion, and all the 6 training/test set partitions. Since the RMSE values have the same scale, mean centring was chosen as preprocessing.

5. Results and discussion

5.1. PLS results

To relate the colour and textural information contained in the RGB images of grape samples with the corresponding amount of total anthocyanins, 198 PLS models were computed on the whole (=54 single block models + 48 low-level data fusion models + 96 mid-level data fusion models) applying the different techniques and approaches described in the previous sections. The results of the corresponding PLS models are reported in Table 2 and expressed as RMSECV and RMSEP values associated to each model. For a more immediate interpretation of the results, the cells are coloured according to the corresponding RMSE values, from green for lowest RMSE values to red for highest RMSE values. For a more detailed visualization of the outcomes, the results of all the 198 PLS models are reported in the Supplementary Material (Tables S1–S6) including also the corresponding number of LVs, RMSEC, and R^2 (calculated in calibration, cross-validation and prediction of external test set).

This section collects some general comments from a qualitative point of view of the results reported in Table 2. A more detailed and quantitative evaluation of the statistically significant differences between the RMSEP values of the different models is presented in Section 5.2.1, while Section 5.2.2 reports an overall evaluation of the RMSECV and RMSEP values by PCA.

Regarding the feature extraction technique, the highest RMSECV and RMSEP values are generally reached when considering as single blocks the GLCM features at all distances and SCTD. As a matter of fact, GLCM principally focuses on local texture features in an image, compressing this information in few statistical descriptors (i.e., contrast, correlation, energy, homogeneity and entropy) calculated over the averaged GLCMs in the four spatial orientations. This technique is very effective in describing local texture properties of regular and repeated patterns, providing a strong data compression [38]. However, GLCM retains too weak information for this specific application, since the images of the berries have irregular patterns and the progressive accumulation of anthocyanins in red grape has a main impact on the colour of berries rather than texture. On the other side, the immediate and overall estimate of the colour and texture properties of an image returned by SCTD does not appear representative enough to account for the intrinsic variability of the berries within the same grape sample/ripening stage and a more detailed image characterisation is required for the issue at hand. This latter aspect could explain the generally better RMSECV and RMSEP values reached when calculating the models considering CLG and TXG as single blocks. Indeed, these two types of signals seem to allow a more thorough representation of the image properties compared to GLCM and SCTD. In particular, CLG signals represent a sort of fingerprint of colour information in the images, enabling the comparison and modelling of colour variation between different images and relate this information with the anthocyanins content. On the other hand, TXG provide a more detailed description of texture properties compared to GLCM and SCTD. Indeed, in TXG texture information is not simply coded using statistical descriptors but the signals are composed by frequency distribution curves simultaneously accounting for spatial variations of pixel intensities differences in the three colour channels and considering different orientations. In addition, it has to be considered that TXG performances could be further optimized by evaluating the effect of the number of PCs and the number of bins on model performances.

For a more comprehensive evaluation of the PLS models obtained from single blocks, the corresponding predictions related to the Original

Table 2

Results of the PLS models. The cells are coloured according to the corresponding RMSE values, from green for lowest RMSE values to red for highest RMSE values.

		Vin A 2020		Vin B 2020		Vin A 2021		Vin B 2021		Vin C 2021		Original	
	Technique	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**
Single Blocks	**CLG**	164	122	160	186	135	195	151	120	174	130	144	158
	TXG D5	159	158	198	204	232	141	201	181	174	209	200	179
	TXG D10	129	222	178	181	191	131	212	159	167	198	175	142
	TXG D15	123	250	165	132	190	116	217	134	153	192	158	126
	GLCM D1	282	306	250	503	305	308	305	266	307	245	256	249
	GLCM D5	268	261	241	357	291	235	219	306	283	230	241	224
	GLCM D10	251	258	246	288	267	376	233	286	254	260	232	194
	GLCM D15	246	250	242	245	244	227	254	275	229	251	223	181
	SCTD	243	259	262	289	265	190	235	243	215	354	234	238
Low-Level Data Fusion	**CLG + TXG D5**	201	140	168	246	140	260	176	160	199	154	163	149
	CLG + TXG D10	192	124	156	218	140	232	187	144	182	136	154	154
	CLG + TXG D15	173	123	159	202	134	209	168	138	164	127	138	141
	CLG + GLCM D1	209	141	167	275	145	283	208	188	207	143	160	189
	CLG + GLCM D5	183	128	146	209	146	260	167	114	189	187	152	178
	CLG + GLCM D10	153	134	147	177	143	208	156	119	172	183	156	156
	CLG + GLCM D15	141	136	149	163	141	172	144	119	159	184	159	139
	CLG + SCTD	143	147	142	134	138	149	134	124	141	161	151	142
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	99	104	111	174	90	158	117	137	100	130	106	148
	CLG + TXG D10	119	145	95	182	80	158	116	127	118	126	99	136
	CLG + TXG D15	113	155	103	154	83	155	101	119	111	122	95	125
	CLG + GLCM D1	184	161	167	295	115	224	159	199	162	141	118	161
	CLG + GLCM D5	149	132	148	216	94	292	126	205	144	130	114	164
	CLG + GLCM D10	119	130	142	177	102	306	110	188	136	121	121	158
	CLG + GLCM D15	115	132	140	157	100	285	108	192	129	124	122	146
	CLG + SCTD	115	141	133	202	124	176	120	140	131	144	123	170
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	121	105	115	201	83	199	132	126	142	125	112	153
	CLG + TXG D10	133	113	103	165	84	229	124	133	124	123	104	158
	CLG + TXG D15	121	113	112	157	81	201	104	127	103	123	93	148
	CLG + GLCM D1	158	139	152	287	110	240	149	154	163	174	120	166
	CLG + GLCM D5	152	106	143	220	97	289	129	175	139	125	120	169
	CLG + GLCM D10	129	124	141	173	93	307	120	159	147	173	122	159
	CLG + GLCM D15	115	132	138	152	102	284	120	159	139	171	124	152
	CLG + SCTD	116	140	134	174	112	174	120	128	130	135	114	157

TR/TS partition are reported in Figs. S3–S6 of Supplementary Material.

When moving from the PLS models calculated on single data blocks to those calculated after data fusion, there seems to be a general improvement of the results, especially when considering mid-level data fusion in cross-validation. However, it has to be mentioned that, interestingly, CLG and TXG as single blocks reach relatively low errors for some training/test set partitions even with respect to data fusion. In addition, the different pretreatments applied for mid-level data fusion seem not to have a notable impact on cross-validation and prediction performances. As already mentioned, the prediction of anthocyanins from RGB images of grape samples is mainly a colour-related problem and CLG approach allows accounting for this aspect. However, the calibration models seem to generally benefit, albeit modestly, from the inclusion of texture features.

With regard to the different partitioning in training and test set, generally lower prediction performances were achieved when considering samples belonging to Vin A 2021 and Vin B 2020 in the test set, but these aspects will be further detailed in the following section.

As an example of the predictions obtained from fused datasets, Fig. S7 of Supplementary Material reports the PLS outcomes of mid-level data fusion of CLG and TXG D15.

5.2. *Statistical evaluation of the results*

5.2.1. *ANOVA*

As described in Section 4.3, ANOVA was performed on the RMSEP values of all the models externally validated with only the samples collected in a specific vineyard and in a specific year, to have a more accurate estimate of the variability possibly due to the different vineyards in each vintage.

The RMSEP values obtained from the models calculated using the single blocks, i.e., considering the performances of the different feature extraction techniques individually, were evaluated using two-way ANOVA without interaction. The results, reported in Table 3, show that significant effects are ascribable only to the different feature extraction techniques. Based on this outcome, the least significant difference (LSD) multiple comparison test ($\alpha = 5\ \%$) was performed to compare each feature extraction method based on the RMSEP values across the 5 training/test set partitionings (Fig. 5). The RMSEP values obtained considering both CLG and TXG (at all the tested distances) were significantly lower with respect to those obtained by the PLS models calculated with GLCM and SCTD. As already outlined on Section 5.1, the best performances of both CLG and TXG suggest that in the problem under investigation these two feature extraction methods allow a better representation of image properties of interest for the problem at hand. Considering CLG and TXG, no statistically significant differences were found in the RMSEP values of the two feature extraction techniques. This result suggests that the TXG approach may represent a hybrid method that can account for both colour and texture features, since the signals are based on the spatial differences of pixel intensities in the three colour channels.

The RMSEP values of the PLS models obtained using data fusion were evaluated using three-way ANOVA with interaction, and the results are reported in Table 4. ANOVA highlighted a statistically significant difference between the different vineyards-years and between the different feature extraction techniques, and also the technique × vineyard-year interaction resulted statistically significant. No statistically significant effects were ascribable to the different data fusion approaches tested in this study. Figs. 6 and 7 report the LSD intervals ($\alpha = 5\ \%$) for each combination of feature extraction techniques and for each vineyard-year partition, respectively.

Table 3
Results of two-way ANOVA on the RMSEP of the models calculated using the single blocks.

Source	Sum of squares	Degrees of freedom	Mean square	F	p- value[a]
Technique	165461.5	8	20682.68	7.323883	**1.7E-05**
Vineyard-year	14496.91	4	3624.228	1.283365	0.297108
Error	90368.15	32	2824.005		
Total	270326.5	44			

[a] Statistically significant differences (p-value <0.05) are highlighted in bold.

According to Fig. 6 the fusion of CLG with TXG calculated considering a distance of 15 pixels led to significantly lower RMSEP values with respect to the models calculated by fusing CLG with GLCM, independently of the pixel distance considered and of the data fusion approach. Interestingly, the RMSEP values of the models calculated by combining CLG and SCTD are not significantly different from those of the models obtained from the fusion of CLG and TXG. Considering that SCTD alone provided significantly higher RMSEP values compared to CLG and TXG as single blocks, we can assume that the improved performances obtained after fusing CLG and SCTD are mainly ascribable to the contribution of CLG block.

Even if the vineyard-year effect resulted not significant in the ANOVA based on the results of single blocks, this effect was significant when considering the ANOVA of data fusion results. As shown in Fig. 7, significantly higher RMSEP values are obtained when considering Vin A 2021 and Vin B 2020 as external test sets. In accordance with the results reported in Table 2, considering the different training/test partitions based on vineyard and vintage year, the lowest prediction performances were obtained for Vin A 2021, especially when considering the data fusion approaches. However, the application of the texturegrams as single technique at all the pixel distances considered seem to reach relatively low RMSEP values when the samples of Vin A 2021 are in the external test set.

The lower performances observed for Vin A 2021 and Vin B 2020 partitions suggest a lack of representativeness of the samples belonging to the other vineyards and/or years to model Vin A 2021 and Vin B 2020 samples. Therefore, a larger dataset, including samples from more vineyards and more vintages, would be required to better cover geographical and time related variability sources.

On the other hand, Vin C 2021 partition was initially supposed to be the most critical case, since the external validation is performed on samples belonging to a vineyard not considered in the training set. However, the outcomes show a general good predictive capability for Vin C 2021 samples, which is not significantly different from Vin A 2020 and Vin B 2021 partitions.

5.2.2. *Exploratory data analysis of the results*

To have an overall comparison of the different model performances and a general overview of the contribution coming from feature extraction techniques, data fusion approaches and training/test set partitions to the characterisation of the models under investigation, an exploratory analysis was performed by means of PCA on the results achieved with both the single techniques and their combinations through data fusion. i.e., on the data reported in Table 2. The dataset used for PCA was therefore composed by a total of 33 modeling approaches (9 single techniques + 8 technique combinations × 3 data fusion approaches) described by the corresponding RMSECV and RMSEP values calculated according to each training/test set partition tested, i.e. by a total of 12 variables. Fig. 8a reports the score plot of the first two PCs of the PCA model, accounting for 87.14 % of total data variance. In Fig. 8a, the different objects (corresponding to the sets of modeling approaches) are coloured according to the image feature extraction technique (both as single block and fused with CLG), while their marker shape refers to the data block approach (i.e., single block, low-level data fusion, mid-level data fusion of autoscaled score vectors, mid-level data fusion of mean centred and block scaled score vectors).

Fig. 8b reports the corresponding PC1 vs. PC2 loading plot, where the variables are coloured according to the cross validation (RMSECV) or prediction error (RMSEP) and labelled according to the respective

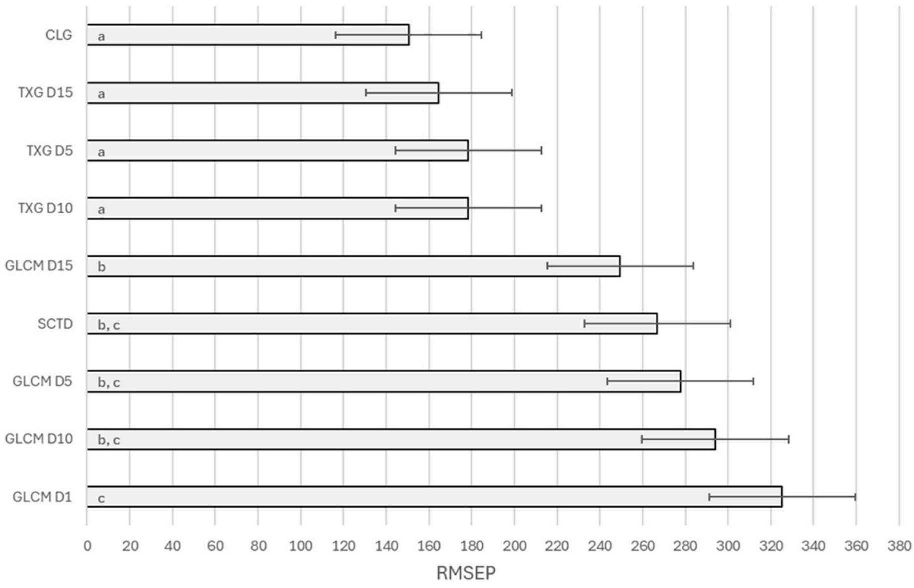


Fig. 5. Means and LSD intervals (α = 5 %) of RMSEP values of the feature extraction techniques resulting from the ANOVA model reported in Table 3. Statistically significant differences are highlighted with different letters.

Table 4
Results of three-way ANOVA with interaction performed on the RMSEP of the models obtained using data fusion approaches.

Source	Sum of squares	Degrees of freedom	Mean square	F	p- value[a]
Technique	37125.3	7	5303.6	8.63	**3.8829E-07**
Data fusion	192.3	2	96.1	0.16	0.8556
Vineyard-year	157502.9	4	39375.7	64.04	**3.0673E-20**
Technique × Data fusion	14698.4	14	1049.9	1.71	0.0800
Technique × Vineyard-year	47289.2	28	1688.9	2.75	**6.4497E-04**
Data fusion × Vineyard-year	9808.9	8	1226.1	1.99	0.0639
Error	34433.1	56	614.9		
Total	301050.1	119			

[a] Statistically significant differences (p-value <0.05) are highlighted in bold.

training/test partition.

It is possible to observe that PC1 mainly describes the general error associated with each technique throughout all the training/test set partitions, since all the elements of PC1 loading vector are positive and with similar values. Therefore, along PC1 lower errors are associated to negative scores, and vice versa.

PC2 instead mainly reflects the differences between the RMSECV and the corresponding RMSEP values obtained for the various PLS models and is particularly influenced by the high RMSEP values generally observed for Vin A 2021 and Vin B 2020.

The models obtained with single blocks are more spread when compared with the ones obtained after data fusion. GLCM and SCTD models as single blocks are located at the highest positive score values of PC1 and distant from the remainder models. CLG on its own is close to the low- and mid-level data fusion approaches. According to the results reported in Table 2, CLG as single technique generally appears to have similar performance to low-level data fusion especially in cross-validation, but also to mid-level data fusion in prediction. At the same time, as shown in Fig. 8a, also TXG approach shows interesting performances when considered individually, at all the pixel distances explored in this study. In accordance with this, Fig. 5 and Table 2 give evidence of the performances obtained when TXG are considered as single

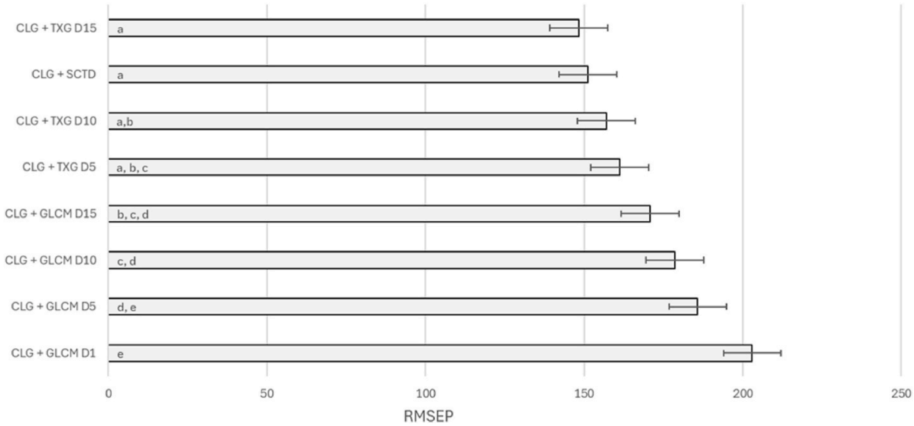


Fig. 6. Means and LSD intervals (α = 5 %) of RMSEP values of the different combinations of feature extraction techniques resulting from the ANOVA model reported in Table 4. Statistically significant differences are highlighted with different letters.

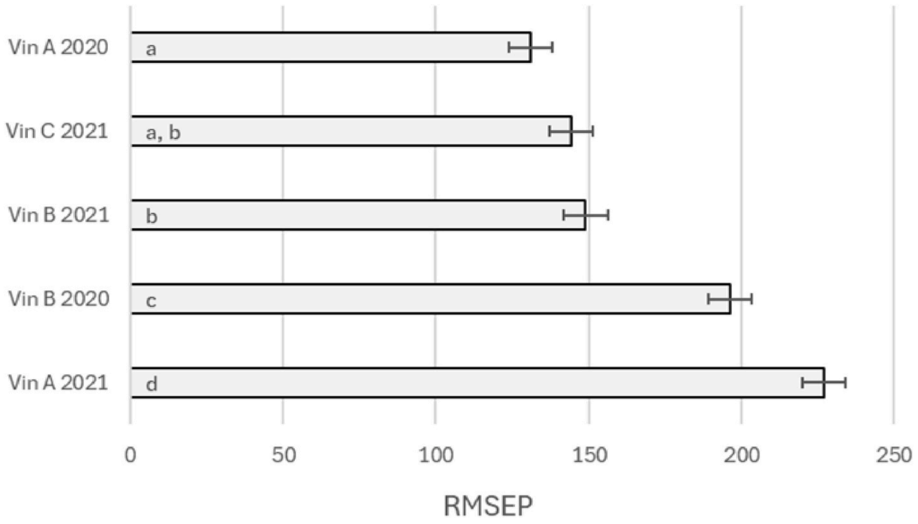


Fig. 7. Means and LSD intervals (α = 5 %) of the RMSEP values of the different vineyard-year partitioning resulting from the ANOVA model reported in Table 4. Statistically significant differences are highlighted with different letters.

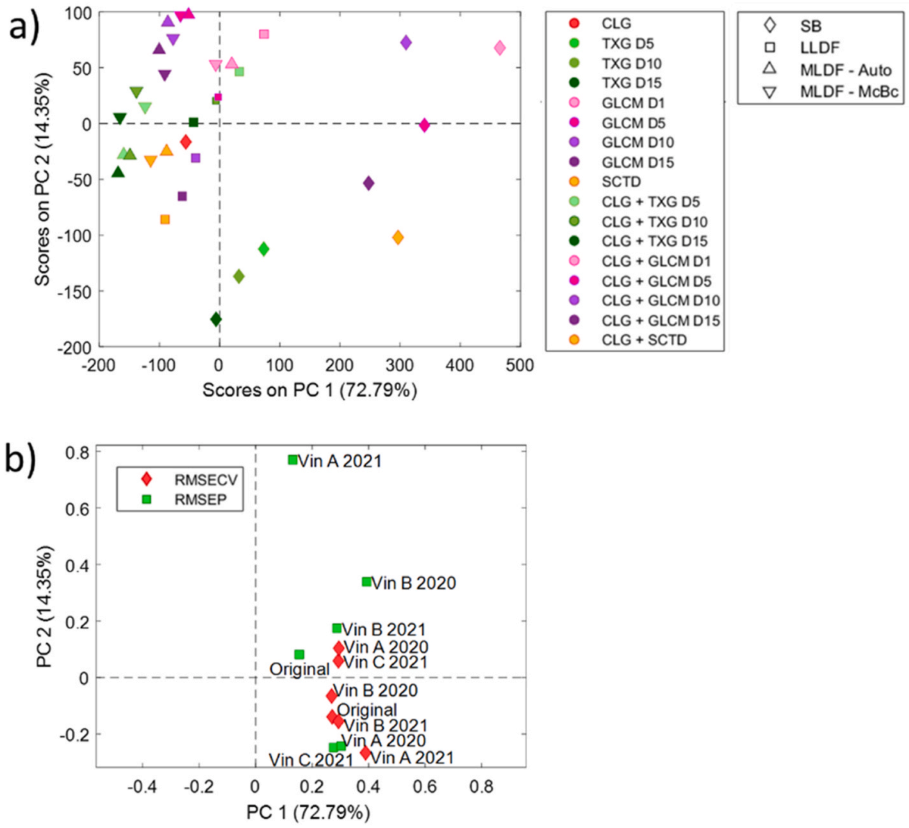


Fig. 8. Results of the PCA model calculated on the data reported in Table 2. In a) PC1 vs. PC2 score plot, where the different models are coloured according to feature extraction technique, while the marker shape represents single block modelling (diamond), low-level data fusion (square), mid-level data fusion with autoscaling (upward-pointing triangle) and mid-level data fusion with mean centering and block scaling (downward-pointing triangle); in b) PC1 and PC2 loading plot where the variables are coloured according to the cross-validation (red) or test set prediction (green) error, while labels refer to the type of training/test set partition. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

technique, which are comparable to those obtained with CLG, and for some training/test set partitions TXG perform even better than the colourgrams. For example, when Vin A 2021 is in the external test set, TXG provided lower RMSEP values when compared with the colourgrams, independently of the pixel distance considered. Indeed, the models related to TXG calculated at a distance of 5, 10 and 15 pixels are located at the lowest score values on PC2 which, according to Fig. 8b, accounts for the difference between RMSEP and the corresponding RMSECV value in particular for Vin A 2021 and, at a minor extent, Vin B 2020 training/test set partitions. According to the score plot in Fig. 8a, the highest differences between RMSEP and RMSECV values to predict these vineyards are associated to the models calculated by fusing at a mid-level the colourgrams with the GLCM at the different distances, by considering the GLCM D1 and GLCM D10 as single blocks and by fusing at a low-level the colourgrams with GLCM D1 and with TXG D5.

The overall higher errors were obtained when considering GLCM D1 alone; in fact, the models related to this single block are located at the extreme positive value of PC1, corresponding to generally high values of both RMSECV and RMSEP, and at high values of PC2, indicating RMSEP values much higher than the (already high) RMSECV ones. This is consistent with the ANOVA results reported in Figs. 5 and 6, were GLCM D1 always shows the worst performance.

As shown in Fig. 8a and in accordance with Table 2, the different preprocessing methods tested for mid-level data fusion do not appear to have a substantial impact on the model performances.

Concerning both GLCM and TXG, a trend along PC1 is observed in Fig. 8a, where increasing pixel distance leads to lower PC1 values, i.e., to generally lower RMSECV and RMSEP values.

With reference to the scores on PC1, Fig. 8a shows that the objects related to mid-level data fusion models are generally characterised by slightly lower PC1 score values, when compared to the corresponding models calculated with low-level data fusion. This suggests as a general trend that mid-level data fusion tends to provide overall best performances, as one might expect. In fact, mid-level data fusion implies a sort of “cleaning” of the useless information from the data. Conversely, low-level data fusion is a more conservative data fusion approach on one side, but it is less effective in selecting only the relevant information for the problem at hand. Comparing these results with those of three-way ANOVA on the RMSEP values reported in Table 4, where no significant differences were observed among the different data fusion approaches, the results of PCA suggest that this trend on PC1 on the one hand is due to the generally better performances in cross-validation of mid-level data fusion models, and on the other hand to the different performances of the considered feature extraction techniques.

With regard to the technique, the two models obtained by mid-level fusion of CLG and TXG D15 show the lowest PC1 scores, suggesting an interesting, even if slight, improvement of the model performances when combining the information brought by the colourgrams and by the texturegrams.

6. Conclusions

In the present study, we discussed the application of different image-level techniques able to extract and analyse colour and texture features of a benchmark dataset of RGB images of red grape samples to build PLS models and relate the image features with anthocyanins content. As a colour-related technique we considered the colourgrams (CLG) approach, while for the extraction of texture features, we used literature methods such as SCTD and GLCM. In addition, we also proposed an alternative method to extract and codify the texture features of an RGB image into a one-dimensional signal, named texturegram (TXG).

The predictive abilities of the different colour and texture techniques were compared. Colourgrams and texturegrams allowed achieving significantly better prediction results compared to GLCM and SCTD, probably due to the more detailed representation of the image properties allowed by these two types of signals. In fact, while the output parameters of both GLCM and SCTD are a limited set of summary statistics, texturegrams consist of a set of histograms describing the distributions of pixel texture-related properties in different orthogonal directions defined by PCA. Another aspect that makes texturegrams completely different from GLCM and SCTD is that a TXG describes the *differences* between pixels at a fixed distance along all the horizontal, vertical and diagonal directions, while GLCM is focused on the *co-occurrences* of the intensity values of couples of pixels at a given distance and angle, and SCTD summarize colour *variability*.

To evaluate the possible benefits of integrating colour and texture information, we also performed data fusion by combining colourgrams with each one of the considered texture methods. As a result, we obtained general slight improvement in model performances. Overall encouraging results were achieved by fusing at the mid-level the colourgrams with the texturegrams, suggesting an interesting, even if slight, improvement of the model performances when combining the information brought by the colourgrams and the texturegrams in a unique solution. As possible future refinement of the models obtained from fused data blocks, multi-block methods such as Multi-Block PLS (MB-PLS), Sequential and Orthogonalized PLS (SO-PLS) and Response-Oriented Sequential Alternation (ROSA) could be applied to improve the prediction results and exploit the complimentary information of colour and texture descriptors [49].

However, considering practical applications, both CLG and TXG as single blocks could represent a good compromise between the predictive capability and computational effort required for the issue explored in this work. In fact, the performances obtained with both approaches as single blocks appear to be comparable to those reached with data fusion, with the advantage of requiring the computation of only one type of feature.

As a matter of fact, the anthocyanins content in grape berries has a notable influence on colour properties rather than on texture ones. The interest in texture features in this work could represent a supplementary improvement of the model performances, by potentially characterizing the images with regard to elements like the background or the presence of berries pedicels.

Based on these considerations, for a more balanced evaluation of the potential of the different texture features and their possible combination with the colour-related ones, future studies are required to be carried out on a more systematically designed dataset. This could also allow not only to analyse more in depth the effects of key parameters like pixel distances (that here was investigated considering only the three distance levels of 5, 10 and 15 pixels), number of histograms bins and number of PCs, but also to better explore the theoretical basis for the advantage of TXG over GLCM and SCTD. Even if we conceived and proposed the texturegrams to be principally focused on texture properties of an image, the outcomes of this study suggest the potential use of this technique as a hybrid tool to simultaneously explore both colour and texture properties. This outcome can be motivated considering that these signals rely on the spatial differences of pixel intensities in the three colour channels.

Furthermore, the algorithm used to calculate TXG could be further optimized considering faster computational methods to calculate the common PCs space like e.g., those based on randomisation [50], making this approach a fast and effective tool to describe colour and texture properties of large datasets of images.

A final consideration regards the potential advantages of using the texturegrams approach in combination with multivariate calibration or classification algorithms, compared with current mainstream Deep Learning techniques. The first advantage of the proposed approach over DL is model interpretability, in order to critically evaluate which information contributes significantly to the estimation of the parameters of interest, In fact, while the interpretation of the most informative texturegram regions (e.g., those corresponding to the higher values of the PLS VIP scores) can be easily obtained by reconstructing in the original domain the corresponding pixels, the evaluation of the contribution of the original features extracted by a DL technique is very difficult or often impossible. The second advantage of TXG over DL is that TXG allow calculating robust PLS calibration models even with a limited number of samples compared with DL methods. Indeed, the correct computation of DL models generally requires a training phase on a huge number of samples, which often is not feasible when dealing with chemical, biological and agricultural problems, where the determination of the reference analytical parameters often requires a considerable effort.

CRediT authorship contribution statement

Camilla Menozzi: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **José Manuel Prats-Montalbán:** Writing – review & editing, Supervision, Software, Methodology, Conceptualization. **Rosalba Calvini:** Writing – review & editing, Software, Methodology, Conceptualization. **Alessandro Ulrici:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

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.

Acknowledgements

Rosalba Calvini would like to thank the Italian funding programme Fondo Sociale Europeo REACT-EU - PON “Ricerca e Innovazione” 2014–2020 – Azione IV.6 Contratti di ricerca su tematiche Green (D.M. 1062 del 10/08/2021) for supporting her research (CUP: E95F21002330001; contract number 17-G-13884–4).

The authors wish also to thank Dr. Giovanni Nigro, Dr. Paola Tessarin and Dr. Domenico Bossio for technical assistance and support during grape harvesting and image acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2025.105446.

Data availability

Data will be made available on request.

References

[1] S. Cubero, M.P. Diago, J. Blasco, J. Tardaguila, J.M. Prats-Montalbán, J. Ibáñez, J. Tello, N. Aleixos, A new method for assessment of bunch compactness using automated image analysis, Aust. J. Grape Wine Res. 21 (2015) 101–109, https://doi.org/10.1111/ajgw.12118.

[2] M. Meenu, C. Kurade, B.C. Neelapu, S. Kalra, H.S. Ramaswamy, Y. Yu, A concise review on food quality assessment using digital image processing, Trends Food Sci. Technol. 118 (2021) 106–124, https://doi.org/10.1016/j.tifs.2021.09.014.

[3] D. Wu, D.-W. Sun, Colour measurements by computer vision for food quality control – a review, Trends Food Sci. Technol. 29 (2013) 5–20, https://doi.org/10.1016/j.tifs.2012.08.004.

[4] C. Zheng, D.-W. Sun, L. Zheng, Recent developments and applications of image features for food quality evaluation and inspection – a review, Trends Food Sci. Technol. 17 (2006) 642–655, https://doi.org/10.1016/j.tifs.2006.06.005.

[5] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. SMC-3 (1973) 610–621, https://doi.org/10.1109/TSMC.1973.4309314.

[6] J.M. Prats-Montalbán, A. De Juan, A. Ferrer, Multivariate image analysis: a review with applications, Chemometr. Intell. Lab. Syst. 107 (2011) 1–23, https://doi.org/10.1016/j.chemolab.2011.03.002.

[7] R. Calvini, G. Foca, A. Ulrici, Data dimensionality reduction and data fusion for fast characterization of green coffee samples using hyperspectral sensors, Anal. Bioanal. Chem. 408 (2016) 7351–7366, https://doi.org/10.1007/s00216-016-9713-7.

[8] R. Archana, P.S.E. Jeevaraj, Deep learning models for digital image processing: a review, Artif. Intell. Rev. 57 (2024) 11, https://doi.org/10.1007/s10462-023-10631-z.

[9] S.M. Javidan, A. Banakar, K. Rahnama, K.A. Vakilian, Y. Ampatzidis, Feature engineering to identify plant diseases using image processing and artificial intelligence: a comprehensive review, Smart Agr. Technol. 8 (2024) 100480, https://doi.org/10.1016/J.ATECH.2024.100480.

[10] S. Günaydın, E. Ropelewska, K. Sacilik, N. Çetin, Exploration of machine learning models based on the image texture of dried carrot slices for classification, J. Food Compos. Anal. 129 (2024) 106063, https://doi.org/10.1016/j.jfca.2024.106063.

[11] S. Verdú, I. García, C. Roda, J.M. Barat, R. Grau, A. Ferrer, J.M. Prats-Montalbán, Multivariate image analysis for assessment of textural attributes in transglutaminase-reconstituted meat, Chemometr. Intell. Lab. Syst. 256 (2025) 105280, https://doi.org/10.1016/j.chemolab.2024.105280.

[12] C. Niu, X. Ying, G. Pei, M. Hu, G. Zhai, Review of the deep learning for food image processing, Int. J. Agric. Biol. Eng. 17 (2024) 15–30, https://doi.org/10.25165/j.ijabe.20241705.8975.

[13] H. Il Yang, S.G. Min, J.H. Yang, J.B. Eun, Y.B. Chung, A novel hybrid-view technique for accurate mass estimation of kimchi cabbage using computer vision, J. Food Eng. 378 (2024) 112126, https://doi.org/10.1016/J.JFOODENG.2024.112126.

[14] P.L. Rolando, R. Calvini, G. Foca, A. Ulrici, Mixture design and multivariate image analysis to monitor the colour of strawberry yoghurt purée, Microchem. J. 194 (2023) 109222, https://doi.org/10.1016/j.microc.2023.109222.

[15] A. Antonelli, M. Cocchi, P. Fava, G. Foca, G.C. Franchini, D. Manzini, A. Ulrici, Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm, Anal. Chim. Acta 515 (2004) 3–13, https://doi.org/10.1016/j.aca.2004.01.005.

[16] A. Biancolillo, R. Ferretti, C. Scappaticci, M. Foschi, A.A. D'Archivio, M. Di Santo, L. Di Martino, Development of a non-destructive tool based on E-eye and agro-morphological descriptors for the characterization and classification of different brassicaceae landraces, Appl. Sci. 13 (2023) 6591, https://doi.org/10.3390/app13116591.

[17] A. Borin, M.F. Ferrão, C. Mello, L. Cordi, L.C.M. Pataca, N. Durán, R.J. Poppi, Quantification of Lactobacillus in fermented milk by multivariate image analysis with least-squares support-vector machines, Anal. Bioanal. Chem. 387 (2007) 1105–1112, https://doi.org/10.1007/s00216-006-0971-7.

[18] J.F. Lopes, A.P.A.C. Barbon, G. Orlandi, R. Calvini, D.P. Lo Fiego, A. Ulrici, S. Barbon, Dual Stage Image Analysis for a complex pattern classification task: ham veining defect detection, Biosyst. Eng. 191 (2020), https://doi.org/10.1016/j.biosystemseng.2020.01.008.

[19] R. Calvini, G. Orlandi, G. Foca, A. Ulrici, Colourgrams GUI: a graphical user-friendly interface for the analysis of large datasets of RGB images, Chemometr. Intell. Lab. Syst. 196 (2020) 103915, https://doi.org/10.1016/j.chemolab.2019.103915.

[20] P.H. Gonçalves Dias Diniz, Chemometrics-assisted color histogram-based analytical systems, J. Chemom. 34 (2020) e3242, https://doi.org/10.1002/cem.3242.

[21] A. Giraudo, R. Calvini, G. Orlandi, A. Ulrici, F. Geobaldo, F. Savorani, Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams, Food Control 94 (2018) 233–240, https://doi.org/10.1016/j.foodcont.2018.07.018.

[22] G. Orlandi, R. Calvini, L. Pigani, G. Foca, G. Vasile Simone, A. Antonelli, A. Ulrici, Electronic eye for the prediction of parameters related to grape ripening, Talanta 186 (2018) 381–388, https://doi.org/10.1016/j.talanta.2018.04.076.

[23] M.H. Bharati, J.J. Liu, J.F. MacGregor, Image texture analysis: methods and comparisons, Chemometr. Intell. Lab. Syst. 72 (2004) 57–71, https://doi.org/10.1016/j.chemolab.2004.02.005.

[24] F. Tomita, S. Tsuji, Statistical texture analysis, in: Computer Analysis of Visual Textures, Springer US, Boston, MA, 1990, pp. 13–36. http://link.springer.com/10.1007/978-1-4613-1553-7_2.

[25] M. Bevk, I. Kononenko, A statistical approach to texture description of medical images: a preliminary study, in: Proceedings of 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), IEEE Computer. Soc, n.d.: pp. 239–244. https://doi.org/10.1109/CBMS.2002.1011383.

[26] I. Kononenko, M. Kukar, Data preprocessing, in: Machine Learning and Data Mining, Elsevier, 2007, pp. 181–211. https://linkinghub.elsevier.com/retrieve/pii/B9781904275213500071.

[27] G.N. Srinivasan, G. Shobha, Statistical texture analysis, Proc. World Acad. Sci. Eng. Technol. 36 (2008) 1264–1269, files/367/Srinivasan - 2008 - Statistical Texture Analysis.pdf.

[28] J.M. Prats-Montalban, F. Lopez, J.M. Valiente, A. Ferrer, Multivariate statistical projection methods to perform robust feature extraction and classification in surface grading, J. Electron. Imag. 17 (2008) 031106, https://doi.org/10.1117/1.2957886.

[29] F. López, J.-M. Valiente, J.-M. Prats, Surface grading using Soft colour-texture descriptors, in: A. Sanfeliu, M.L. Cortés (Eds.), Progress in Pattern Recognition, Image Analysis and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 13–23. http://link.springer.com/10.1007/11578079_2.

[30] N. Aggarwal, R.K. Agrawal, First and second order statistics features for classification of magnetic resonance brain images, J. Signal Inf. Process. 3 (2012) 146–153, https://doi.org/10.4236/jsip.2012.32019.

[31] A.A. Chandini, U. Maheswari B, Improved quality detection technique for fruits using GLCM and MultiClass SVM, in: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2018, pp. 150–155, https://doi.org/10.1109/ICACCI.2018.8554876.

[32] C. Malegori, L. Franzetti, R. Guidetti, E. Casiraghi, R. Rossi, GLCM, an image analysis technique for early detection of biofilm, J. Food Eng. 185 (2016) 48–55, https://doi.org/10.1016/j.jfoodeng.2016.04.001.

[33] J. Paliwal, N.S. Visen, D.S. Jayas, N.D.G. White, Cereal grain and dockage identification using machine vision, Biosyst. Eng. 85 (2003) 51–57, https://doi.org/10.1016/S1537-5110(03)00034-5.

[34] M.H. Bharati, J.F. MacGregor, Texture analysis of images using principal component analysis, in: H. McCann, D.M. Scott (Eds.), Intell Syst Smart Manuf, 2001, p. 27, https://doi.org/10.1117/12.417179.

[35] J.M. Prats-Montalbán, A. Ferrer, Integration of colour and textural information in multivariate image analysis: defect detection and classification issues, J. Chemom. 21 (2007) 10–23, https://doi.org/10.1002/CEM.1026.

[36] L. Pieszczek, M. Daszykowski, Integrating hyperspectrograms with class modeling techniques for the construction of an effective expert system: quality control of

pharmaceutical tablets based on near-infrared hyperspectral imaging, J. Pharm. Biomed. Anal. (2025) 116697, https://doi.org/10.1016/j.jpba.2025.116697.

[37] C. Menozzi, R. Calvini, G. Nigro, P. Tessarin, D. Bossio, M. Calderisi, V. Ferrari, G. Foca, A. Ulrici, Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes, Microchem. J. 191 (2023) 108811, https://doi.org/10.1016/j.microc.2023.108811.

[38] I. Lubis, R. Siregar, Recognition of medicinal plant leaf patterns using morphology-based and GLCM feature extraction, J. Artif. Intellig. Eng. Appl. (JAIEA) 3 (2024) 834–839, https://doi.org/10.59934/jaiea.v3i3.522.

[39] M. Gasparro, A.R. Caputo, C. Bergamini, P. Crupi, M.F. Cardone, R. Perniola, D. Antonacci, Sangiovese and its offspring in southern Italy, Mol. Biotechnol. 54 (2013) 581–589, https://doi.org/10.1007/s12033-012-9600-1.

[40] Y. Glories, Maturité phénolique du raisin, conséquences technologiques: application aux millésimes 1991 et, 1992, in: 1993, pp. 56–61.

[41] N. Saint Cricq de Gaulejac, N. Vivas, Y. Glories, Maturation phénolique des raisines rouges. Relation avec la qualité des vins. Comparison des cépages Merlot et Tempranillo, Progrès Agricole et Viticole 115 (1998) 306–318.

[42] P. Ribéreau-Gayon, Le dosage des anthocyanes dans le vins rouge, Bull. Soc. Chim. 9 (1965) 2649, files/269/1573668924051904640.html.

[43] S. Río Segade, E. Soto Vázquez, E. Díaz Losada, Influence of ripeness grade on accumulation and extractability of grape skin anthocyanins in different cultivars, J. Food Compos. Anal. 21 (2008) 599–607, https://doi.org/10.1016/j.jfca.2008.04.006.

[44] S. Sommer, S.D. Cohen, Comparison of different extraction methods to predict anthocyanin concentration and color characteristics of red wines, Fermentation 4 (2018) 39, https://doi.org/10.3390/fermentation4020039.

[45] J.M. Prats-Montalbán, A. Ferrer, R. Bro, T. Hancewicz, Prediction of skin quality properties by different Multivariate Image Analysis methodologies, Chemometr. Intell. Lab. Syst. 96 (2009) 6–13, https://doi.org/10.1016/j.chemolab.2008.10.012.

[46] D.A. Clausi, An analysis of co-occurrence texture statistics as a function of grey level quantization, Can. J. Rem. Sens. 28 (2002) 45–62, https://doi.org/10.5589/m02-004.

[47] R. Calvini, L. Pigani, Toward the development of combined artificial sensing systems for food quality evaluation: a review on the application of data fusion of electronic noses, electronic tongues and electronic eyes, Sensors 22 (2022) 577, https://doi.org/10.3390/S22020577, 22 (2022) 577.

[48] G. Orlandi, R. Calvini, G. Foca, L. Pigani, G. Vasile Simone, A. Ulrici, Data fusion of electronic eye and electronic tongue signals to monitor grape ripening, Talanta 195 (2019) 181–189, https://doi.org/10.1016/j.talanta.2018.11.046.

[49] P. Mishra, J.-M. Roger, D. Jouan-Rimbaud-Bouveresse, A. Biancolillo, F. Marini, A. Nordon, D.N. Rutledge, Recent trends in multi-block data analysis in chemometrics for multi-source data integration, TrAC, Trends Anal. Chem. 137 (2021) 116206, https://doi.org/10.1016/j.trac.2021.116206.

[50] J.P. Cruz-Tirado, J.M. Amigo, D.F. Barbin, S. Kucheryavskiy, Data reduction by randomization subsampling for the study of large hyperspectral datasets, Anal. Chim. Acta 1209 (2022) 339793, https://doi.org/10.1016/j.aca.2022.339793.