On-the-go table grape ripeness estimation via proximal snapshot hyperspectral imaging

Riccardo Bertoglio [a],[*], Manuel Piliego [a], Paolo Guadagna [b], Matteo Gatti [b], Stefano Poni [b], Matteo Matteucci [a]

[a] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via Ponzio 34/5, Milan, 20133, Italy
[b] Department of Sustainable Crop Production, Università Cattolica del Sacro Cuore, Via Emilia Parmense, 84, Piacenza, 29122, Italy

ARTICLE INFO

ABSTRACT

The monitoring of grapes for ripeness estimation is a practice that enables fruit harvesting at the optimal time. Hyperspectral Imaging (HSI) represents a non-destructive and high-throughput alternative to traditional laboratory analyses. Current literature approaches perform hyperspectral measurements using line scan sensors or low-resolution static snapshot cameras, which hinder a fast per-bunch ripeness characterization. We propose a framework for on-the-go collection and processing of proximal snapshot hyperspectral images to estimate single bunch ripeness parameters. Focusing on table grapes (*Vitis vinifera* L. cv. Red Globe), we collected images under natural illumination with a hyperspectral camera (500–900 nm) mounted on a moving vehicle in an experimental block sited in Piacenza, Italy. We investigated images collected in August and September 2021 representing two ripening stages. The composition of the imaged grape bunches was determined through laboratory chemical analyses to predict Total Soluble Solids (TSS) and anthocyanin concentration. The images were pre-processed via multimodal image registration to correct the unalignment of bands due to the vehicle motion, and the single bunches were automatically identified on false RGB images through a Mask Region-Convolutional Neural Network (Mask R-CNN) instance segmentation network. The mean spectra of the bunches were used as input features of a Partial Least Squares Regression (PLSR) model to predict the chemical parameters at single bunch and whole vine scales. The regression model of TSS had an R^2 (10-fold nested cross-validation) of 0.75 and 0.85 on a per-bunch and per-vine basis, respectively. The regression model of anthocyanin had an R^2 of 0.68 and 0.49 on a per-bunch and per-vine basis, respectively. The results suggest the potential of using snapshot hyperspectral images for high-throughput analysis of a per-bunch grape ripeness estimation. The method described in this study could give valuable information to improve grape ripening monitoring and management of harvest operations and even allow for precise and automated robotic harvesting.

1. Introduction

Monitoring grape quality is of high importance to determine the ripening stage and the optimal harvest time for both wine and table grapes (Poni et al., 2018). Grapes are non-climacteric fruits, so they do not ripe any further after harvest (Prasanna et al., 2007). Moreover, ripening is influenced by many factors, including environmental conditions, grape variety, soil type, and growing techniques (Meléndez et al., 2013; Dai et al., 2011; CooMbe and McCarthy, 2000; Gouot et al., 2019; Poni et al., 2023). Grape ripening is usually described as the accumulation of sugars and is measured in terms of TSS; however, other chemical parameters are taken into account to determine optimal ripeness (Meléndez et al., 2013), such as phenolics, volatile

compounds, titratable acidity, and organic acids. Phenolic compounds, like anthocyanins, are mostly accumulated in berry skin and seeds, having a critical impact on the sensory attributes of grapes, influencing color and astringency (Cheynier et al., 2006). Conventional approaches to determining the grape quality and ripeness rely upon destructive measurements of key chemical components concentration determined on representative berry samples collected in vineyards.

However, destructive approaches are laborious and can only be applied to a limited number of samples. A typical sampling protocol consists of at least 4–5 grape samples, each made of at least 100 berries taken from different bunches and positions within the same bunch. Berries should be intact to avoid juice loss, hence biasing the

---

subsequently measured berry fresh weight. Along with the inherent time-consuming approach, seasonal repeated sampling from a given vine can also affect and modify the natural ripening pattern as a progressively decreasing pending crop will presumably mature faster. Thus, there is a need for a shift towards high-throughput automated solutions (Ye et al., 2022; Vrochidou et al., 2021). The application of modern technologies, including robotics and non-destructive techniques like HSI, enables fast monitoring. HSI gathers information on how electromagnetic radiations of various wavelengths interact with the matter (ElMasry and Sun, 2010). The capability of this technology for extracting information from the target is notable since the relative contribution of each light wavelength interaction depends on the sample's chemical composition and physical parameters. Initially limited to controlled environments such as laboratories (Gabrielli et al., 2021; Gomes et al., 2021; Gao and hua Xu, 2022; Xu et al., 2023), it has gradually begun to be used directly in the field thanks to miniaturization and improved computing and storage capabilities of the measuring devices (Power et al., 2019; Vrochidou et al., 2021; Lu et al., 2020). Grape quality attributes can be monitored directly in the vineyard, with the acquisition process handled remotely by airborne (Sousa et al., 2022; Matese et al., 2022) or ground vehicles (Gutiérrez et al., 2019; Benelli et al., 2021), producing proximal hyperspectral images with high spatial resolution. HSI sensors can be broadly based on their acquisition modes, whether pushbroom or snapshot (Adão et al., 2017). Pushbroom cameras acquire a single column of pixels for all spectral bands, while snapshot cameras capture the entire pixel frame either in a single shot or multiple shots for each band sequentially. It is worth noting that multiple-shot cameras typically offer higher resolution compared to single-shot cameras.

Although some in-field studies have shown positive results in predicting chemical quality parameters through HSI, most studies perform measurements using line scan pushbroom technology which hinders the identification of single grape bunches. Gutiérrez et al. (2019) performed in-field and on-the-go acquisitions using a pushbroom Resonon Pika L VNIR HSI camera installed on an all-terrain vehicle. They scanned 180 plants of a Tempranillo cultivar, measuring 300 bands from 400 nm to 1000 nm with 300 pixels of spatial resolution. Epsilon-Support Vector Machines with Gaussian kernel were used for predicting TSS and anthocyanins concentration, respectively obtaining an R^2 of 0.92, 0.83 on the test set. Finally, TSS and anthocyanin concentration spatial–temporal prediction maps were generated using multilevel b-spline interpolation. Fernández-Novales et al. (2019), in a very similar setting, performed on-the-go acquisitions by using a contactless Visible-Short Wave Near Infrared PSS spectrometer, measuring 215 datapoints per spectrum in the 570–990 nm range. The grape clusters' absorbance spectra were used in PLSR for predicting TSS (test set $R^2 = 0.95$), anthocyanins ($R^2 = 0.79$), and total polyphenols ($R^2 = 0.43$). Kalopesa et al. (2023) observed TSS values in Chardonnay, Malagouzia, Sauvignon-Blanc, and Syrah varieties over harvest and pre-harvest stages. Hyperspectral measurements were performed using PSR+3500 spectrometer, a highly accurate contact probe spectrometer covering the entire Visible Near Infrared and Short-Wave Infrared spectrum (350–2500 nm), directly applied to the berry. PLSR, Random Forest, SVR, and 1D CNN as predictive methods for each distinct variety were compared, with 1D CNN prevailing over the alternatives for three varieties out of four. Benelli et al. (2021) monitored Soluble Solids Content in a Sangiovese vineyard over thirteen days during pre-harvest and harvest time. The hyperspectral acquisitions were performed via a NanoHyperspec VNIR pushbroom camera with a spectral resolution of 272 bands between 400 nm and 1000 nm. The exposure time was adjusted by calibrating with a white high-reflectance matter panel at the same distance as the vineyard row. The reflectance spectra of the sections were used as input for a PLSR, yielding a cross-validation (CV) R^2 of 0.768 and an RMSECV of 0.79 °Brix.

Studies performing in-field acquisitions with snapshot cameras are just a few and they mostly perform static measurements, hindering a high-throughput analysis. Rodríguez-Pulido et al. (2022) studied the sugar concentration in single Tempranillo and Syrah berries. They acquired hyperspectral images using a Specim IQ hyperspectral camera (range 400–1000 nm). It is a handheld pushbroom camera that outputs 512 lines of 512 pixels, thus simulating a snapshot camera (Behmann et al., 2018). However, its line acquisition modality limits its use to static image collection. They acquired bunch images under direct sunlight twice a week from July to harvest (mid-August) in 2020. Multiple white calibration measurements were made throughout each session, employing a certified reflectance device to correct for daylight changes. Using the laboratory data as the calibration set and in-field data as the test set, a PLSR model achieved an R^2 of 0.88 when used on the reflectance spectra and an R^2 of 0.91 on the absorbance spectra. The varying illumination conditions during in-field experiments are a significant obstacle that restricts the generalization of hyperspectral-based prediction models. Tsakiridis et al. (2023) proposed denoising autoencoders for passing from the raw hyperspectral signals to the corresponding reflectance signals to account for the changing light conditions. Considering different pretreatment options, the authors trained deep fully connected (DAE) and deep convolutional autoencoders (DCAE) using the raw hyperspectral signals as input and reference reflectance signals as target. The resulting spectra were used to compare the predictive capabilities for TSS of Extreme Gradient Boosting (XGBoost) Random Forest, Cubist, and SVR over the single varieties, with the best model's test RMSE ranging between 1.66 °Brix and 2.29 °Brix.

In summary, the majority of research has concentrated on acquiring data in controlled environments, which often lack the transferability needed for real-world outdoor applications. While pushbroom technology has been prevalent in experiments conducted in uncontrolled outdoor settings, it struggles with identifying individual grape bunches, presenting limitations. Only a few studies have investigated the use of snapshot cameras for in-field applications, albeit primarily with static imaging approaches. Consequently, there exists a significant gap in research addressing on-the-go acquisitions with snapshot cameras, which enables high-throughput analysis with precise per-bunch accuracy. Indeed, in-field studies mainly focus on predicting grape parameters in bunch clusters (comprising bunches from multiple vines) or selected single grape berries. The direct investigation of single bunches would allow a characterization of the chemical variability in the vineyard with higher spatial resolution, enlarging and refining the amount of information that can be leveraged for targeted cultivation and harvest procedures, e.g., location-specific irrigation and nurturing and automated picking of single bunches. Finally, there is a lack of studies that investigate hyperspectral sensing for table grapes (Gabrielli et al., 2021).

The objective of this study is to propose a pipeline for acquiring and analyzing on-the-go proximal snapshot hyperspectral images aimed at predicting Total Soluble Solids (TSS) and anthocyanin concentration. Hyperspectral images were obtained using a camera mounted on a moving vehicle, enabling high-throughput data collection. However, the movement of the camera caused spectral band misalignment, prompting an investigation into multimodal techniques to effectively align the band images. To the best of our knowledge, this is the first study to explore on-the-go proximal snapshot hyperspectral sensing. Leveraging snapshot images allowed us to utilize an instance segmentation model to identify individual grape bunches, a crucial step for automating table grape harvesting, which is currently a manual and labor-intensive process.

2. Materials and methods

The elaboration pipeline comprises three main stages, as illustrated in Fig. 1. The first stage involves on-the-go acquisition of hyperspectral images of bunches held by table grape vines. At each collection moment, samples are taken from the plants and subsequently analyzed
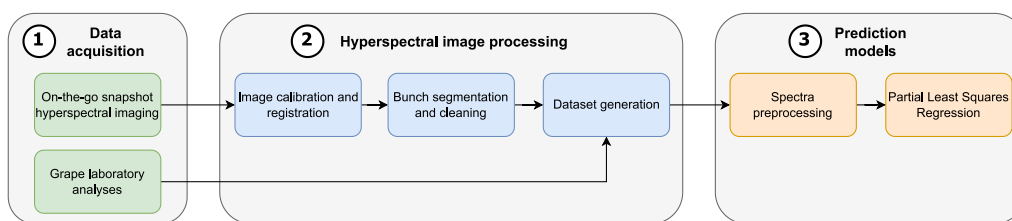
Fig. 1. Study elaboration pipeline consisting of three stages: data acquisition, hyperspectral image processing, and the development of prediction models.

in the laboratory to measure the chemical composition of bunches. The second stage involves processing these images to prepare them for the prediction phase. Initial steps include calibration and registration. Following this, bunches are segmented and cleaned, with mean spectra computed for each bunch, forming the dataset. The final stage entails spectra preprocessing, serving as input to a PLSR model for predicting chemical parameters. The code used in our analysis is available at the following link: https://github.com/AIRLab-POLIMI/hs-ripeness-estimation.

2.1. Image acquisitions

The experimental site was a vineyard facility of the Università Cattolica del Sacro Cuore, located in Piacenza, Italy. *Vitis vinifera* L. cultivar Red Globe grapevines were grown outdoors in 251 pots arranged in two NE-SW oriented rows and grafted on Selection Oppenheim 4 (SO4) rootstock. Each vine was vertically shoot-positioned and hedgerow-trained with a 10–11 nodes fruiting cane rising about 80 cm from the ground. A set of three surmounting catch wires was used to facilitate hedgerow training, resulting in a $1.3\,\mathrm{m}$ tall canopy wall. On the East-facing side of each row, three bunches per vine were designated to correlate the measured chemical parameters with the corresponding hyperspectral data. To promote a fully defoliated fruit zone and reduce the occurrence of berry sunburn (Gatti et al., 2015), the leaves surrounding grape bunches were gradually removed between fruit set (BBCH 71) and berry touch (BBCH 79) (Lorenz et al., 1995). To address excessive fruit density, eight vines were subjected to crop thinning before veraison; accordingly only one bunch per shoot was maintained to mitigate fruit occlusion as well as to enhance fruit ripening (Tang et al., 2023; Gatti et al., 2012). At the same time, three bunches per vine were tagged from shoots located on the basal, median and distal nodes of the fruiting cane, to be analyzed at harvest. The remaining untagged fruits were then used for pre-harvest acquisitions and berry collection as described below.

Hyperspectral acquisitions were conducted in 2021 at two ripening stages on August 23 (pre-harvest), and September 6 (harvest). The hyperspectral cubes were acquired in situ using a Senop HSC-2 hyperspectral camera. The camera was mounted on a SCOUT 2.0 AgileX robotic platform (see Fig. 2). This four-wheeled differential driving robot was teleoperated to navigate the vineyard while maintaining a centered trajectory between the vineyard rows. Each hypercube comprised 49 bands uniformly distributed within the 500–900 nm range, supplemented by 3 bands with specific wavelengths to generate false RGB images (R: 600 nm, G: 556.5 nm, B: 510 nm), resulting in a total of 52 bands. The resolution of each band image was 1024×1024 pixels. The camera was oriented with its principal axis perpendicular to the plant rows.

2.2. Chemical analyses

On each date, once the hyperspectral measurements were acquired, grape samples per each vine were randomly collected and brought to the laboratory to determine the TSS and total anthocyanin concentration. In August 2021, two exposed berries were randomly chosen from the two sides (external and internal) of three unlabeled bunches,

resulting in a twelve-berry sample per each of the vines considered within the study (14 vines in August and 16 vines in September). Six of these were immediately processed to determine the TSS must concentration, while the remaining six berries were stored in a freezer at $-18\,°\mathrm{C}$ for subsequent quantification of total anthocyanins. Thus, measurements were expressed on a per-vine basis on this date. At harvest in September 2021, all the labeled fruits for a total of 58 bunches were individually analyzed. Accordingly, fourteen berries (seven on each side of the bunch) were reserved for total anthocyanin analysis, while the remaining grapes were immediately crushed to determine TSS concentration. At each sampling time, the fresh weight of the berries reserved for the anthocyanin assessment was measured. Fig. 3 illustrates boxplots depicting the chemical parameters for both dates and both sets of analyses (per-bunch and per-vine). From the per-vine values, it is evident that both TSS and anthocyanin levels have higher quartiles in September 2021 compared to August 2021, indicating a clear sign of riper grapes.

Fresh grapes were crushed, and the resulting juice was processed to determine TSS (°Brix) concentration by using a temperature-compensated digital refractometer (ATAGO DBX-55, Tokyo, Japan). Total anthocyanins were assessed according to Iland et al. (2011). Frozen berries were initially halved with a scalpel and then homogenized for 1 min at 20,000 rpm with an Ultra-Turrax homogenizer (Rose Scientific Ltd., Edmonton, Canada). A portion of the homogenate was then transferred to a pre-tared centrifuge tube, enriched with 10 mL of aqueous ethanol (50%, pH 5.0), and periodically mixed for 1 h before centrifugation at $959 \times g$ for 5 min. Then, 0.5 mL of the extract was added to 10 mL of 1M HCl (1M solution of hydrochloric acid), mixed, and left to stand for 3 h. Finally, absorbance was measured at $520\,\mathrm{nm}$ using a JascoV-530 UV spectrophotometer (Jasco Analytical Instruments, Easton, MD, USA). The total anthocyanin concentration was expressed in mg per g of fresh berry mass.

2.3. Dataset generation

The process of handling hyperspectral images, which resulted in the creation of the final dataset, is depicted in Fig. 4. Initially, the hyperspectral images underwent geometric and radiometric calibration. Subsequently, the images were registered, and instance segmentation was applied to identify the bunches. Finally, potential inaccuracies were addressed through outlier removal, and the mean spectral signals of the bunches were computed.

To rectify the images for geometrical distortions, we computed intrinsic parameters and distortion coefficients individually for each wavelength. Multiple perspectives of a checkerboard panel with a known square length and size were captured, and these images were utilized to determine the calibration parameters. We implemented the code for the geometric calibration in Python, employing the OpenCV[1] library. We also conducted a radiometric calibration to accommodate varying light conditions on the two collection dates. At the beginning of each data collection session, we imaged a calibration panel, specifically,

---

[1] https://opencv.org.

Fig. 2. The robot while acquiring images in the vineyard with the hyperspectral camera mounted on the side. The camera is highlighted with a red circle.

(a) Per-bunch values.
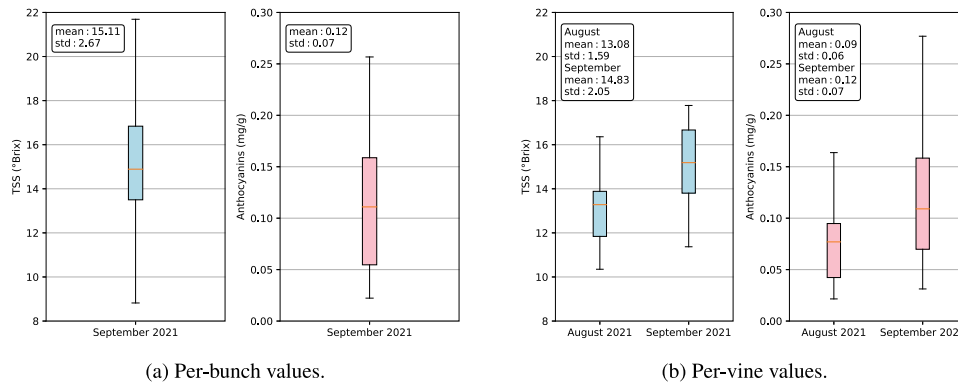
(b) Per-vine values.

Fig. 3. Boxplots of the TSS and total anthocyanins concentration in juice and grapes, respectively depending on single bunch and whole plant measurements. The means and standard deviations (std) of the chemical values are reported in the top-left boxes.
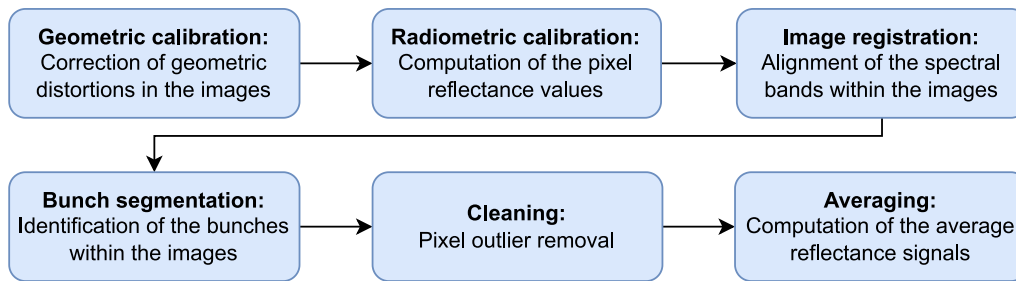

**Fig. 4.** Hyperspectral image processing and dataset generation pipeline.

a Sphere Optics Zenith Lite Diffuse Target SG3515 with a diffuse reflectance of 95%. For the radiometric calibration process, we employed the Linear Regression Method, an empirical model, to directly convert Digital Numbers (the image pixel values) to reflectance (Smith and Milton, 1999). The reflectance for each band can be obtained with the equation:

$$\text{reflectance}_\lambda = \text{offset}_\lambda + \text{gain}_\lambda \text{DN}_\lambda. \tag{1}$$

Here, DN denotes the Digital Number, and offset and gain are the model parameters to be determined. Such an equation is established for every spectral band within the data cube. For each band, the gain and offset are estimated using the mean DN value of the calibration panel, with the reflectance set to the known calibration panel value of 95%. Subsequently, the derived gain and offset values are employed to

calculate reflectance values for plant images by inserting the DN of the grapevine images into the equation.

The movement of the robotic platform and the sequential acquisition of bands resulted in shifted bands in the spatial dimension within the hypercubes. Fig. 5 showcases a pair of color images from a single unregistered hypercube: a false RGB image and a false-color image displaying three extreme bands (the first, middle, and last). It is possible to observe band misalignment, particularly near the borders of objects. Consequently, it became necessary to align the bands comprising the hypercubes by applying image registration techniques. Image registration involves determining a transformation that aligns a moving image with a fixed image. In our study, we performed band-to-reference image registration, designating the band with a wavelength of 704.1 nm as the reference image for each hyperspectral band image. All other images
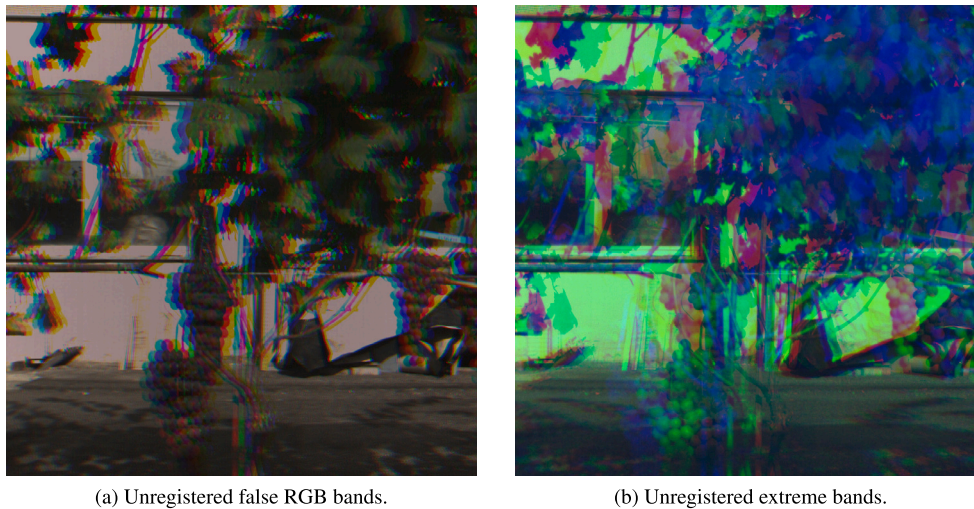
(a) Unregistered false RGB bands.

(b) Unregistered extreme bands.

Fig. 5. Unregistered false RGB (left) and false-color (right) hyperspectral images (September 2021).
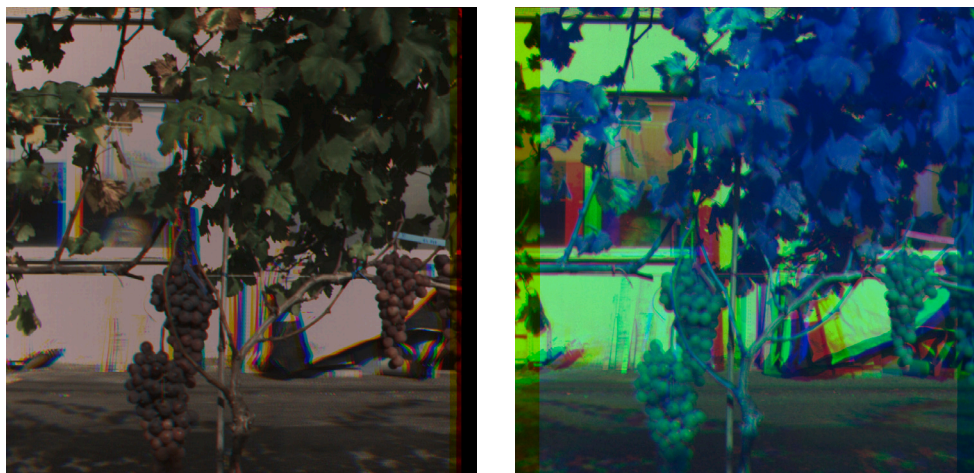
(a) Registered false RGB bands.

(b) Registered extreme bands.

Fig. 6. Registered false RGB (left) and false-color (right) hyperspectral images (September 2021).

at different wavelengths served as moving images to be registered. The choice of the reference band was based on its proximity to the midpoint of the spectral range, located at 700.0 nm. To address our hyperspectral setting, we utilized a multimodal image registration technique (Zitová and Flusser, 2003). Multimodal (or intermodal) registration consists of matching images of the same subject from different modalities; in our case, it involves different acquisition wavelengths.

For image registration, we employed the SimpleITK library (Beare et al., 2018; Yaniv et al., 2018; Lowekamp et al., 2013) through its Python interface. The optimal transformation was determined through the minimization of a similarity metric. Given the multimodal context, we opted for Mutual Information (MI) computed using the method proposed by Mattes et al. (2001). The image registration algorithm requires an initial estimate of the transformation. Consequently, we conducted an initial grid search across a range of plausible translations using the Exhaustive Optimizer. Subsequently, we selected the most suitable translation based on MI and utilized it as the initialization for the registration algorithm. The band-to-reference registration was then executed employing a gradient descent optimizer, a linear interpolator, and a 3-level pyramid multiresolution image representation. Fig. 6 displays a pair of registered images corresponding to the unregistered images in Fig. 5.

After registering the images, we employed an instance segmentation algorithm to discern the pixels corresponding to grape bunches from

the canopy and background. Using an instance segmentation model was made possible by the full-frame capability of the snapshot cameras employed during data acquisitions, distinguishing our approach from other literature methods. Specifically, we utilized a Mask R-CNN model that had been previously trained on RGB images of the same vines. The model's training procedure has been extensively detailed in another article (Chiatti et al., 2023). Initially, a Mask R-CNN model underwent pre-training on a large-scale grape bunch dataset and was subsequently fine-tuned using the VINEyard Piacenza Image Collections (VINEPICs) grape image archive (Bertoglio et al., 2023). The fine-tuned model was applied to false RGB images created by selecting three specific bands (600.0 nm, 556.5 nm, 510.0 nm) from each hypercube. The segmentation masks of instances with a confidence level exceeding 0.99 were retained and saved in COCO 1.0 annotation format[2] for subsequent processing. We underscore that the Mask R-CNN model was applied on false RGB images just for inference. It is important to note that the Mask R-CNN model was applied solely for inference on false RGB images. Notably, without any adaptation, acceptable segmentation of grapes was achieved. Fig. 7 displays two examples of instance segmentation predictions on false RGB images. One example
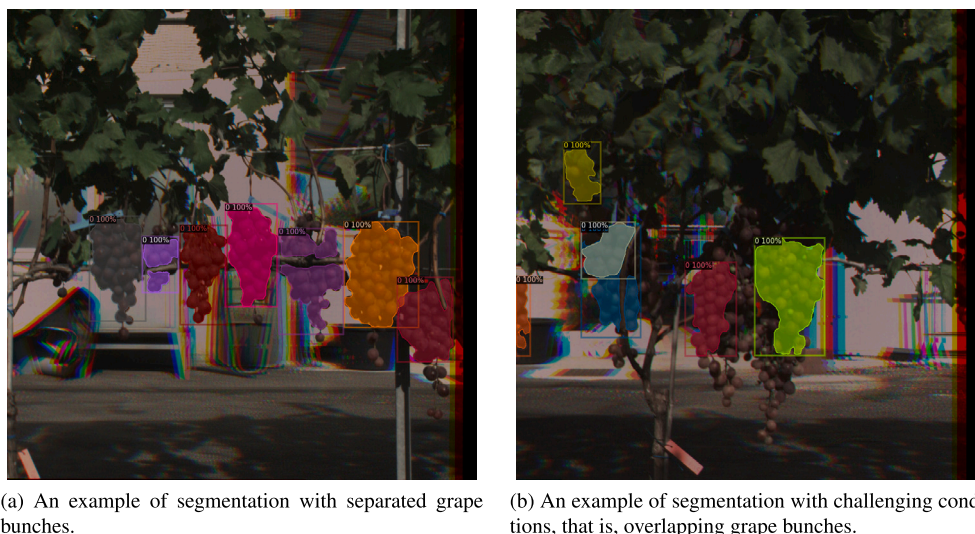
---

[2] https://cocodataset.org.

(a) An example of segmentation with separated grape bunches.

(b) An example of segmentation with challenging conditions, that is, overlapping grape bunches.

Fig. 7. Visualization of the Mask R-CNN inferences on a false RGB images (September 2021).
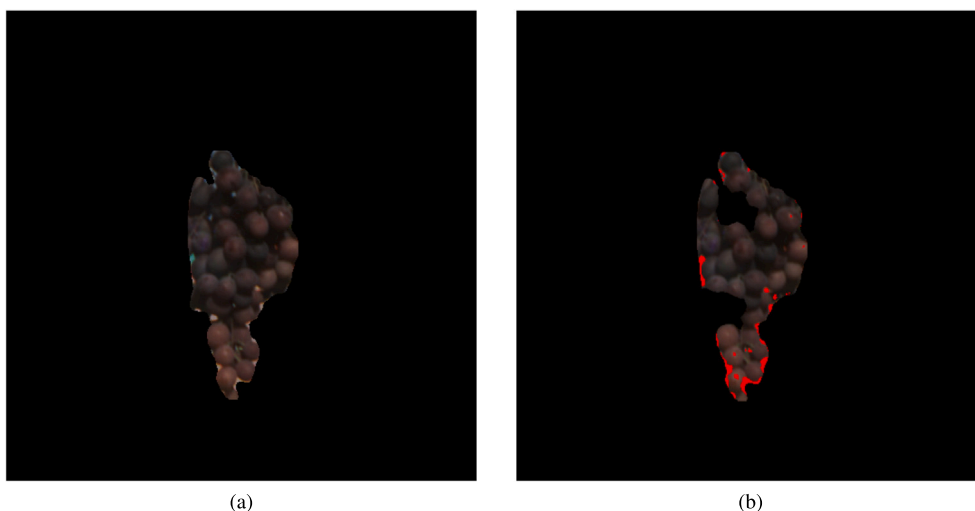

(a)

(b)

Fig. 8. A single bunch segmentation mask applied to a false RGB image before (a) and after (b) IQR cleaning (September 2021). The red pixels are those flagged as outliers and thus ignored.

showcases well-separated grapes where the segmentation performed effectively, while the other presents a more challenging scenario with overlapping grapes. It is worth noting that in some images the instance segmentation model missed certain grape bunches or failed to distinguish individual bunches in situations with overlapping bunches. However, this difficulty is typical in grape bunch detection tasks and is not exclusive to hyperspectral data. Indeed, even for a human labeler, distinguishing single bunches in cluttered clusters would pose challenges. In terms of our analysis scope, all grape bunches under investigation were correctly recognized by the instance segmentation model.

The segmented polygons generated by the Mask R-CNN were manually matched with the corresponding chemical measures. Subsequently, these masks were applied to each band of the 3D hypercubes, isolating grape pixels at every wavelength. We employed an outlier detection method to identify and exclude pixels associated with non-grape targets that were introduced during the registration and segmentation processes. Given the non-Gaussian distribution of reflectance values, we operated within a non-parametric framework, utilizing the InterQuantile Range (IQR) as the statistical dispersion measure. Tukey's

fences (Tukey et al., 1977) were then employed to flag outliers, considering data points lying outside the interval:

$$[\hat{x} - c \cdot IQR, \hat{x} + c \cdot IQR] \quad,$$

where $\hat{x}$ denotes the median, and the empirically determined constant $c = 2.2$ was chosen based on visual inspection. For each 3D hypercube, we applied the IQR detection method individually for each band. Regarding the spatial dimension, all pixels displaying anomalous behavior in at least one spectral band were flagged as outliers and removed from the segmented image (see Fig. 8). Subsequently, we recovered the mean reflectance signals from the relevant pixels in the images. For the analysis focused on individual bunches, the spatial average of reflectance values was computed for the cleaned segmentation mask of the specific bunch. By spatial average, we refer to averaging the reflectance values of grape bunch pixels for each band separately. Conversely, for the analysis focused on vines, the average was computed across all masks of tagged bunches for the plant. The per-bunch dataset comprised 58 data points, representing the 58 grape bunches analyzed during the September 2021 harvest. Meanwhile, the per-vine dataset included 30 data points, consisting of 14 vines examined in August 2021 and 16 vines examined in September 2021.
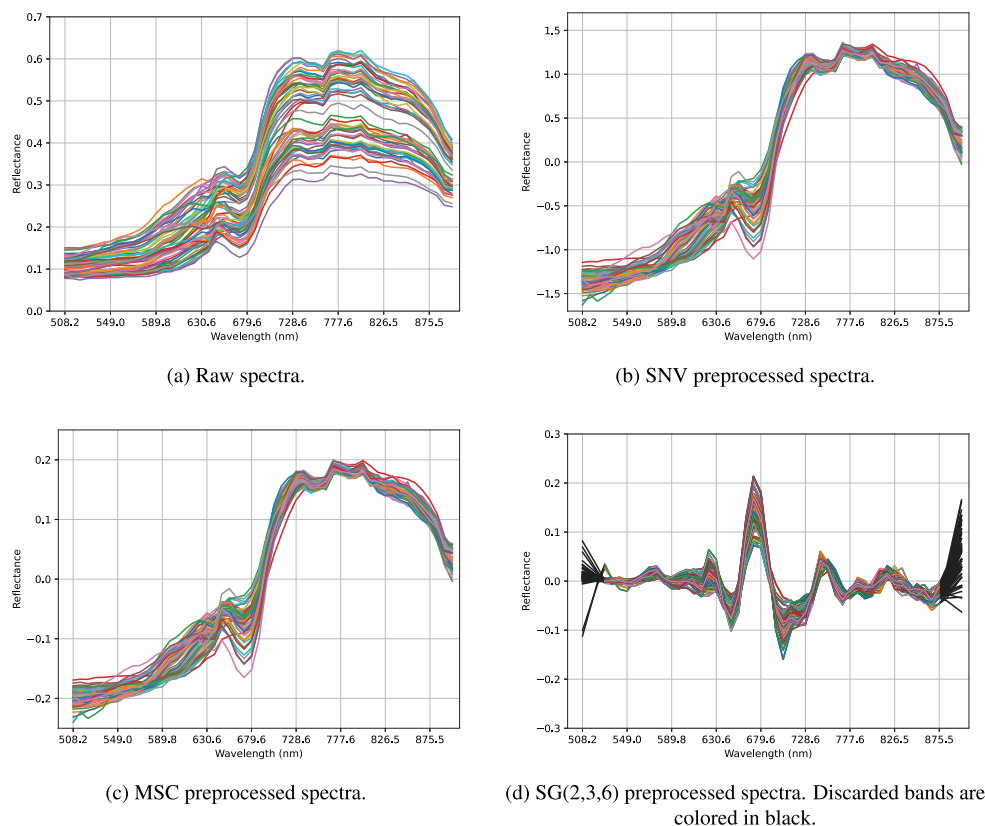
(a) Raw spectra.

(b) SNV preprocessed spectra.

(c) MSC preprocessed spectra.

(d) SG(2,3,6) preprocessed spectra. Discarded bands are colored in black.

Fig. 9. Raw spectra (a) of single bunches for images collected in September 2021, and SNV (b), MSC (c), SG(derivative order, polynomial degree, window length) (d) preprocessed spectra.

2.4. Development of the prediction models

The variation in reflectance spectra stems from three primary sources (Barnes et al., 1989): nonspecific scatter of radiation at the particle surface, variations in spectral path length through the sample, and the sample's chemical composition. The combination of these effects masks the underlying chemical information and constitutes a significant obstacle to interpreting reflectance spectra. These effects exhibit both additive and multiplicative characteristics and vary from one sample to another (Dhanoa et al., 1994). To correct for these effects, we evaluated three commonly used preprocessing methods: Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), and Savitzky–Golay (SG) smoothing and derivative treatments. Fig. 9 displays the raw and preprocessed spectra of the bunches dataset related to the images collected in September 2021.

SNV (Barnes et al., 1989) is a mathematical transformation independent of the dataset, applied individually to each spectrum. SNV normalizes the spectra by subtracting each spectrum's mean and dividing it by its standard deviation. On the other hand, MSC (Isaksson and Næs, 1988) is a dataset-dependent normalization technique designed to mitigate the dominant influence of scatter effects in spectral variation. The fundamental concept is to equalize the scatter level in different samples to the scatter of an ideal reference spectrum.

SG digital filtering (Savitzky and Golay, 1964) is effective for noise reduction while preserving the essential features of the data, that is, the shape and height of waveform peaks and higher moments in the signal. It is commonly coupled with derivative preprocessing. Substituting the original data with their derivatives can be seen as enforcing a high-pass filter and frequency-dependent scaling, de-emphasizing low-frequency trends (Brown et al., 2000). In our study, we performed SG smoothing

and derivative computation using the Python SciPy library.[3] Prior to applying SG treatments, the data underwent initial preprocessing with SNV to enhance the comparability of values. Thus, in the following, we will refer to SG as the combined application of SNV and SG smoothing and derivative treatments. Regarding the SG derivative method, the resulting values at the ends of the spectral range were highly unstable and poorly informative due to the substantial approximation. Consequently, we truncated the spectral signal, excluding the $\lfloor w/2 \rfloor$ most extreme wavelengths at each end, where w indicates the window length of the digital filter.

PLSR methods are a class of algorithms characterized by a common structure for modeling relationships between sets of observed values through a latent representation. PLSR is specifically designed to handle data exhibiting a high degree of collinearity and potentially a larger number of independent variables than samples, a common scenario in spectroscopic analyses. In our study, we employed preprocessed spectra to predict the chemical parameters of interest and assessed the prediction error using Mean Squared Error (MSE). This analysis was conducted by implementing PLSR with the Python scikit-learn library, which serves as an implementation of the PLS1 algorithm.[4]

To robustly evaluate the prediction error of our model and simultaneously compare the three preprocessing methods, we employed a nested CV approach. In the inner loop, we selected the hyperparameter of the PLSR model—specifically, the number of components denoted as l. The outer loop was dedicated to testing the model's performance and comparing the preprocessing methods. The dataset was partitioned into k_{out} outer folds, and for each outer fold, an inner k_{in}-fold CV loop

---

[3] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html

[4] https://scikit-learn.org/stable/modules/generated/sklearn.cross_decomposition.PLSRegression.html

Table 1

Root Mean Squared Error of prediction ($RMSE$), percentage error w.r.t. the target mean ($RMSPE$) and R^2 values of the PLSR models for predicting TSS and anthocyanin content based on the per-bunch dataset using SNV, MSC and SG(derivative order, polynomial degree, window length) preprocessing. l^* is the best value for the number of PLSR components optimized by a single 10-fold CV.

Parameter	Preprocessing	RMSE	RMSPE	R²	l^*
TSS	SNV	1.5824 °Brix	10.4%	0.5910	4
	MSC	1.6378 °Brix	10.8%	0.5802	5
	SG(1,7,8)	1.2655 °Brix	8.4%	0.7475	1
Anthocyanins	SNV	0.0357 mg/g	30.9%	0.6331	3
	MSC	0.0358 mg/g	31.0%	0.6258	3
	SG(1,9,14)	0.0333 mg/g	28.8%	0.6750	1

Table 2

Root Mean Squared Error of prediction ($RMSE$), percentage error w.r.t. the target mean ($RMSPE$) and R^2 values of the PLSR models for predicting TSS and anthocyanin content based on the per-vine dataset using SNV, MSC and SG(derivative order, polynomial degree, window length) preprocessing. l^* is the best value for the number of PLSR components optimized by a single 10-fold CV.

Parameter	Preprocessing	RMSE	RMSPE	R²	l^*
TSS	SNV	0.9629 °Brix	6.9%	0.7290	4
	MSC	0.8846 °Brix	6.3%	0.7697	5
	SG(2,4,15)	0.7294 °Brix	5.2%	0.8524	4
Anthocyanins	SNV	0.0421 mg/g	40.4%	0.4386	2
	MSC	0.0424 mg/g	40.6%	0.4289	2
	SG(2,3,13)	0.0401 mg/g	38.4%	0.4896	2

was employed for model selection based on the remaining $k_{out} - 1$ folds. The optimal model was then retrained on the remaining $k_{out} - 1$ folds and evaluated on the outer fold. The final estimate of the generalization error was computed as the average of the errors across the outer folds. Folds were also stratified based on the target variable, ensuring that each fold adequately represented the variability of the target variable. With $k_{out} = k_{in} = 10$, the number of PLSR components $l_{in,i} \in \{1, \ldots, 15\}$, and the set of preprocessing options $\mathcal{P} \in \{\text{SNV, MSC, SG}(d, p, w)\}$ with $d \in \{0, 1, 2\}$, $p \in \{2, \ldots, 10\}$, $w \in \{2, \ldots, 15\}$ representing the derivative order, polynomial degree, and window length, respectively, we implemented the nested CV as outlined in Algorithm 1. The algorithm output provided the best preprocessing option denoted by P^*—the one with the lowest CV MSE. Subsequently, for each preprocessing method, we determined the optimal number of components l^* using a single 10-fold CV loop on the entire dataset $\mathcal{D}$.

ALGORITHM 1

Nested CV. $\mathcal{P}$ is the set of different preprocessing options, $\mathcal{D}$ is our (per-bunch or per-vine) dataset, $l_{in,i}^*$ is the optimal number of components found within an inner loop.

1: **for** $P \in \mathcal{P}$ **do**
2:     Apply the preprocessing P to the data $\mathcal{D}$
3:     Partition $\mathcal{D}$ in $\{I_1, \ldots, I_{10}\}$ disjoint outer folds
4:     **for** $i = 1, \ldots, 10$ **do**
5:         Select $l_{in,i}^*$ using 10-fold CV on $\mathcal{D} \backslash I_i$
6:         Train a PLSR model of rank $l_{in,i}^*$ on $\mathcal{D} \backslash I_i$
7:         Compute $MSE(l_{in,i}^*; I_i)$
8:     **end for**
9:     Compute $MSECV(P; \mathcal{D}) = \frac{1}{10} \sum_{i=1}^{10} MSE(l_{in,i}^*; I_i)$
10: **end for**
11: Select $P^* = argmin_P\{MSECV(P; \mathcal{D})\}$ as the best preprocessing option

3. Results and discussion

We present the results of predicting TSS (°Brix) and anthocyanin (mg/g) concentration. The nested CV MSE is presented for the three preprocessing methods under scrutiny, namely, SNV, MSC, and SG, and for the two kinds of datasets, that is, per-bunch and per-vine spectra. The per-bunch results are detailed in Table 1.

With the dataset pertaining to bunches, SG outperformed SNV and MSC for both TSS and anthocyanin concentration. In both analyses, the optimal preprocessing involved a 1st derivative treatment. However, the level of smoothing differed between the two cases. For TSS, the optimal hyperparameters for SG were a polynomial of order seven over a window length of eight points (Fig. 10(a)). Conversely, for anthocyanins, the optimal hyperparameters were a polynomial of length nine over a window of length fourteen, resulting in an overall higher level of smoothing (Fig. 10(b)). The PLSR models achieved a coefficient of determination (R^2) of 0.75 for TSS and 0.68 for anthocyanins. When

identifying a single set of optimal hyperparameters through a single-loop CV, the optimal number of components was determined to be one for both analyses.

Considering the per-vine dataset, an SG 2nd derivative treatment achieved better results than the SNV and MSC alternatives for both TSS and anthocyanins (Table 2), with a minimal difference between the three methods in the case of anthocyanins. As in the bunches' analysis, the level of smoothing was more considerable for anthocyanins, using a polynomial of order three and a window length of 13 (Fig. 11(b)), than for TSS, employing a polynomial of order four and a window length of fifteen (Fig. 11(a)). The PLSR model achieved an R^2 of 0.85 for TSS and 0.49 for anthocyanins. Considering the complexity of the PLSR models, a single loop CV indicated four LVs for TSS and two LVs for anthocyanins. SG treatments achieved some level of separation of the reflectance spectra based on the target variables but less than in the per-bunch analysis, leading to more complex modeling choices.

TSS predictions were more accurate for both bunches and vines than those for anthocyanins, aligning with findings in the literature (Ye et al., 2022; Gutiérrez et al., 2019). This discrepancy may stem from the nonuniform distribution of the chemical components. TSS uniformly develops in grapes, making surface reflectance measurements effective indicators of the entire sample's TSS concentration. However, this rationale only partially applies to anthocyanins. Anthocyanin accumulation during ripening is influenced by sunlight exposure, concentrating on the exposed skin portions of berries. Consequently, anthocyanin concentrations vary within bunches, with higher levels in sun-exposed berries and lower levels in shaded or covered berries. This inherent variability may create a discrepancy between measured reflectance signals, focusing on the external bunch surface, and reference chemical quantities measured in the lab, reflecting the whole berries.

TSS values were better predicted using plants, while the opposite was true for anthocyanins. To interpret this difference in performance, two facts had to be considered: the nature of the recorded data and the different sizes of the two datasets. The reflectance and chemical measurements of the bunches were performed under the same environmental conditions on the same day. They represented grapes at a similar stage of ripening (at least at a temporal level), resulting in reflectance signals that were similar in shape. On the other hand, the plants' dataset comprised measurements from two different dates (August and September 2021). The shape of the reflectance signals changed to a moderate degree between the two acquisition dates (Fig. 12). Studying the differences in the chemical parameters via nonparametric permutation tests (Pesarin and Salmaso, 2010) (which were needed due to the non-gaussianity of the anthocyanins measurements) showed that the TSS content varied significantly between the two dates (p-value: 0.0367). In contrast, the changes in anthocyanin content were less relevant (p-value: 0.2042). A possible interpretation of the higher performance of the plants' dataset for predicting TSS could be associated with this difference: the PLSR model exploited the shape dissimilarities of the signals, which carried information about the underlying target variable, to make its guesses more precise. On the
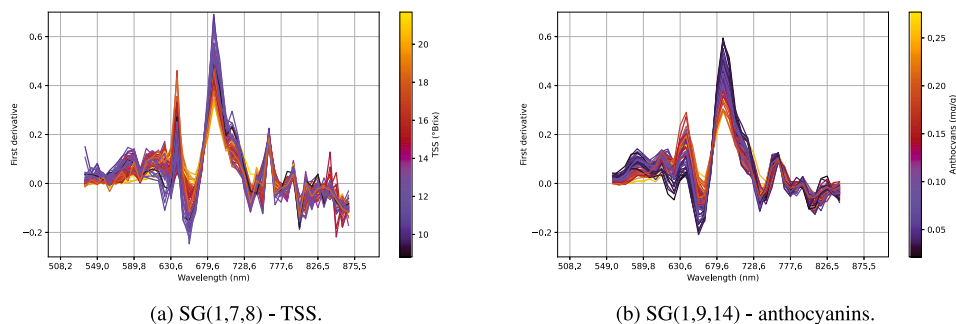
(a) SG(1,7,8) - TSS.

(b) SG(1,9,14) - anthocyanins.

Fig. 10. Per-bunch spectra preprocessed with the optimal pretreatment for TSS and anthocyanins prediction, that is, SG(derivative order, polynomial degree, window length).

(a) SG(2,4,15) - TSS.

(b) SG(2,3,13) - anthocyanins.

Fig. 11. Per-vine spectra preprocessed with the optimal pretreatment for TSS and anthocyanins prediction, that is, SG(derivative order, polynomial degree, window length).
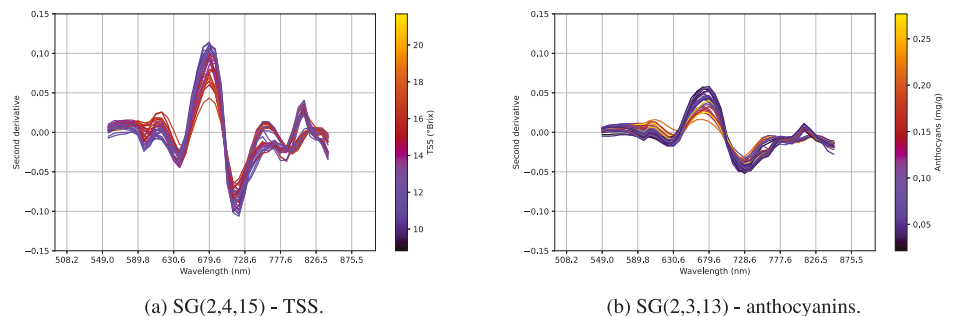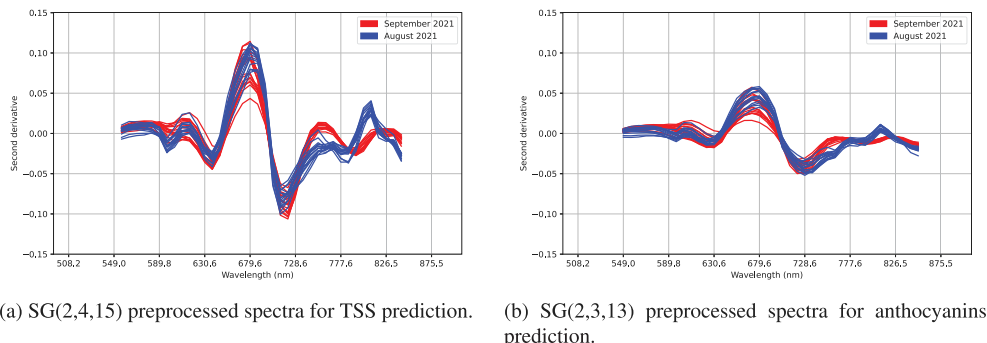

(a) SG(2,4,15) preprocessed spectra for TSS prediction.

(b) SG(2,3,13) preprocessed spectra for anthocyanins prediction.

Fig. 12. Per-vine spectra preprocessed with the optimal pretreatments for TSS and anthocyanins prediction. The spectra are colored based on the acquisition date: red for September 2021 and blue for August 2021.

contrary, the shape dissimilarities were not helpful for the prediction of anthocyanins and possibly masked other informative characteristics in the spectra, effectively lowering the predictive ability of the model.

TSS values exhibited superior predictability when using per-vine data, whereas the reverse was observed for anthocyanins. To comprehend this performance discrepancy, we need to consider the nature of the recorded data. Per-bunch reflectance and chemical measurements were conducted under identical environmental conditions on the same day, representing grapes at a comparable ripening stage, at least temporally. This homogeneity resulted in reflectance signals with similar shapes. Conversely, the per-vine dataset incorporated measurements from two different dates (August and September 2021), causing an increased dissimilarity in the reflectance signal shapes (Fig. 12). Analyzing differences in chemical parameters through nonparametric permutation tests (Pesarin and Salmaso, 2010) (necessary due to the non-Gaussian nature of anthocyanins measurements) revealed significant variation in TSS content between the two dates (p-value: 0.0367). In contrast, changes in anthocyanin content were less relevant (p-value: 0.2042). A plausible interpretation of the superior performance with the per-plant dataset in predicting TSS could be attributed to this distinction: the PLSR model exploited the more pronounced dissimilarities

in signal shapes, which carried information about the underlying target variable, enhancing the precision of its predictions. Conversely, these shape dissimilarities were less pronounced with the anthocyanins and did not contribute significantly to their prediction.

By observing the regression plots for the bunches' dataset in Fig. 13 we noticed how the performance of the models, especially in the case of anthocyanins, was deteriorated by a few problematic data points that were poorly predicted. Implementing outlier detection and removal, a common practice in the literature (Hernández-Hierro et al., 2013; Fernández-Novales et al., 2019; Xu et al., 2023), is expected to yield an improvement in the results. Overall, the performance in predicting TSS and anthocyanins was comparable to similar studies (Benelli et al., 2021; Gutiérrez et al., 2019), except for the prediction of anthocyanins in the per-vine dataset. However, the key characteristic of this study, using a hyperspectral snapshot camera on a moving vehicle, introduces new challenges not currently addressed by other works, such as the need for a registration procedure, which adds an additional source of noise to the data. Nevertheless, snapshot and on-the-go hyperspectral images offer the advantage of providing high-throughput and precise monitoring of individual grape bunches.
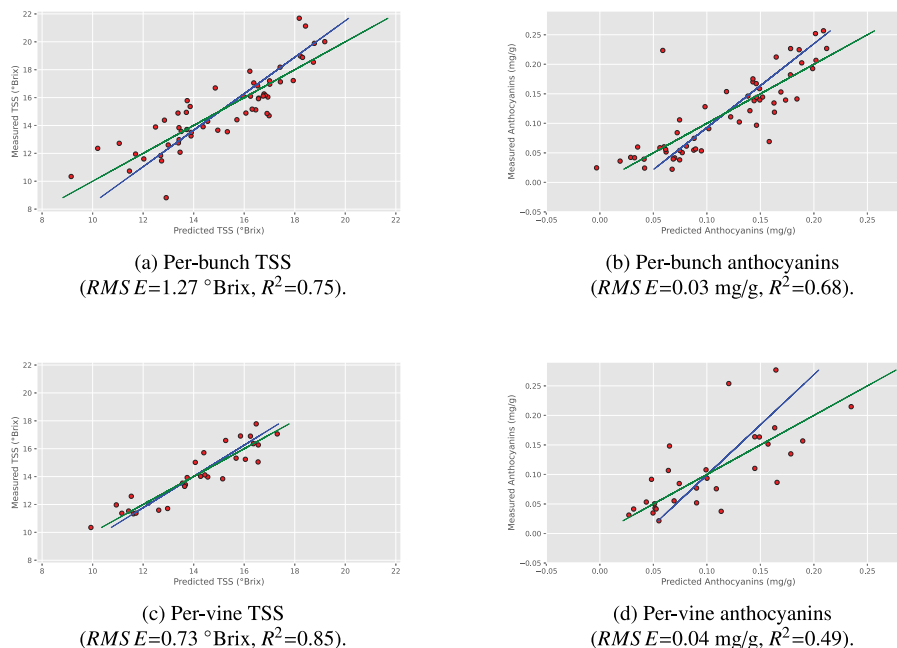
(a) Per-bunch TSS
($RMSE$=1.27 °Brix, R^2=0.75).

(b) Per-bunch anthocyanins
($RMSE$=0.03 mg/g, R^2=0.68).

(c) Per-vine TSS
($RMSE$=0.73 °Brix, R^2=0.85).

(d) Per-vine anthocyanins
($RMSE$=0.04 mg/g, R^2=0.49).

Fig. 13. Regression plots depicting the true vs. predicted nested CV values for TSS and anthocyanin concentrations. The 1:1 line is shown in blue, and the regression line of the samples is presented in green.

4. Conclusions

We presented a pipeline for processing on-the-go snapshot hyperspectral images of grapes in the vineyard, aiming to predict TSS and anthocyanin concentration on a per-bunch and per-vine scale. Due to vehicle movement and the sequential band acquisition, hypercube bands were misaligned, prompting the introduction of a registration technique. Utilizing the full frame of the snapshot camera employed in our study allowed us to employ an instance segmentation model for detecting individual grape bunches. Notably, a model trained solely on RGB images demonstrated strong performance on false RGB hyperspectral registered images, which visually differ from RGB images. We computed the mean pixel value of grape bunches for each band to obtain spectra, and compared three spectra preprocessing techniques, with SG treatments proving superior in both per-bunch and per-vine analyses over SNV and MSC. Successful predictions of chemical indicators were achieved at the bunch and vine level with a PLSR model. Overall, TSS predictions outperformed anthocyanins. This study represents the first pipeline description for acquiring and processing on-the-go proximal snapshot hyperspectral images of grapes. The results highlight HSI as a viable, non-destructive alternative for TSS and anthocyanin analyses. The single bunch identification can enhance the characterization of chemical variability within and between grapevine plants. Potential applications include generating prescription maps to support selective harvesting and enabling automated solutions to the labor-intensive manual harvesting of table grapes. In particular, accurate ripeness information for individual bunches would enable fully automated harvesting by a robot equipped with a gripper arm for picking grape bunches precisely. A strength of our work lies in the simplified spectral representation, with a reduced number (52) of wavelengths compared to other studies in the literature. This results in lower memory requirements, more frequent image acquisitions, and an overall acceleration of the processing pipeline.

A limitation of our study is the small quantity and variety of samples, a consequence of the expensive and time-consuming nature of chemical laboratory analyses. To enhance the robustness of our findings, the validation of predictive models should involve a more extensive dataset and incorporate an independent test set, possibly of a different nature. Future research should focus on testing the generalization capability of the prediction model on unseen vintages and varieties. It is worth noting that in our experimental setup, vine plants were defoliated. In cultivations where this practice is not applied, image collection may face challenges due to foliage occlusions, necessitating active vision or tools like a blower to move leaves. Additionally, predictive analyses should extend to other relevant chemical parameters, such as flavonols, tannins, titratable acidity, pH, and potassium, which are other important indicators of grape ripening. In our study, we conducted inference directly using an instance segmentation model trained on RGB images of the same vineyard. While acceptable segmentation results were achieved, the accuracy should be increased by labeling false RGB hyperspectral images and fine-tuning the model accordingly. Moreover, a larger sample size would enable the effective use of more data-driven predictive methods. In particular, given the rising interest in 1D CNN regression networks for their promising predictive and generalization capabilities, additional exploration is needed. Finally, using snapshot hyperspectral images paves the way for exploring 3D CNN regression models to process masked hypercubes directly, although the associated higher computational costs need careful consideration.

CRediT authorship contribution statement

Riccardo Bertoglio: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Manuel Piliego:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Paolo Guadagna:** Methodology, Data curation, Conceptualization. **Matteo Gatti:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Conceptualization. **Stefano Poni:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization. **Matteo Matteucci:** Writing – review & editing, Supervision, Resources, Methodology, 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.

Data availability

Data will be made available on request.

References

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., Sousa, J.J., 2017. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. Remote Sens. 9 (11), http://dx.doi.org/10.3390/rs9111110.

Barnes, R., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43 (5), 772–777. http://dx.doi.org/10.1366/0003702894202201.

Beare, R., Lowekamp, B., Yaniv, Z., 2018. Image segmentation, registration and characterization in R with SimpleITK. J. Stat. Softw. 86 (8), 1–35. http://dx.doi.org/10.18637/jss.v086.i08.

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M.T., Jussila, J., Salo, H., et al., 2018. Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. Sensors 18 (2), 441. http://dx.doi.org/10.3390/s18020441.

Benelli, A., Cevoli, C., Ragni, L., Fabbri, A., 2021. In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging. Biosyst. Eng. 207, 59–67. http://dx.doi.org/10.1016/j.biosystemseng.2021.04.006.

Bertoglio, R., Gatti, M., Poni, S., Matteucci, M., 2023. VINEyard piacenza image collections - VINEPICs. http://dx.doi.org/10.5281/zenodo.7866442.

Brown, C.D., Vega-Montoto, L., Wentzell, P.D., 2000. Derivative preprocessing and optimal corrections for baseline drift in multivariate calibration. Appl. Spectrosc. 54 (7), 1055–1068.

Cheynier, V., Duenas-Paton, M., Salas, E., Maury, C., Souquet, J.-M., Sarni-Manchado, P., Fulcrand, H., 2006. Structure and properties of wine pigments and tannins. Am. J. Enology Vitic. 57 (3), 298–305. http://dx.doi.org/10.5344/ajev.2006.57.3.298.

Chiatti, A., Bertoglio, R., Catalano, N., Gatti, M., Matteucci, M., 2023. Surgical fine-tuning for grape bunch segmentation under visual domain shifts. In: 2023 European Conference on Mobile Robots. ECMR, IEEE, pp. 1–7. http://dx.doi.org/10.1109/ECMR59166.2023.10256348.

CooMbe, B.G., McCarthy, M., 2000. Dynamics of grape berry growth and physiology of ripening. Aust. J. Grape Wine Res. 6 (2), 131–135. http://dx.doi.org/10.1111/j.1755-0238.2000.tb00171.x.

Dai, Z.W., Ollat, N., Gomès, E., Decroocq, S., Tandonnet, J.-P., Bordenave, L., Pieri, P., Hilbert, G., Kappel, C., van Leeuwen, C., et al., 2011. Ecophysiological, genetic, and molecular causes of variation in grape berry weight and composition: a review. Am. J. Enology Vitic. 62 (4), 413–425. http://dx.doi.org/10.5344/ajev.2011.10116.

Dhanoa, M., Lister, S., Sanderson, R., Barnes, R., 1994. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. J. Near Infrared Spectrosc. 2 (1), 43–47. http://dx.doi.org/10.1255/jnirs.30.

ElMasry, G., Sun, D.-W., 2010. Principles of hyperspectral imaging technology. In: Hyperspectral Imaging for Food Quality Analysis and Control. Elsevier, pp. 3–43. http://dx.doi.org/10.1016/B978-0-12-374753-2.10001-2.

Fernández-Novales, J., Tardáguila, J., Gutiérrez, S., Paz Diago, M., 2019. On-the-go VIS+SW-NIR spectroscopy as a reliable monitoring tool for grape composition within the vineyard. Molecules 24 (15), http://dx.doi.org/10.3390/molecules24152795.

Gabrielli, M., Lançon-Verdier, V., Picouet, P., Maury, C., 2021. Hyperspectral imaging to characterize table grapes. Chemosensors 9 (4), 71. http://dx.doi.org/10.3390/chemosensors9040071.

Gao, S., hua Xu, J., 2022. Hyperspectral image information fusion-based detection of soluble solids content in red globe grapes. Comput. Electron. Agric. 196, 106822. http://dx.doi.org/10.1016/j.compag.2022.106822.

Gatti, M., Bernizzoni, F., Civardi, S., Poni, S., 2012. Effects of cluster thinning and preflowering leaf removal on growth and grape composition in cv. Sangiovese. Am. J. Enology Vitic. 63 (3), 325–332. http://dx.doi.org/10.5344/ajev.2012.11118.

Gatti, M., Garavani, A., Cantatore, A., Parisi, M., Bobeica, N., Merli, M., Vercesi, A., Poni, S., 2015. Interactions of summer pruning techniques and vine performance in the white Vitis vinifera cv. Ortrugo. Aust. J. Grape Wine Res. 21 (1), 80–89. http://dx.doi.org/10.1111/ajgw.12107.

Gomes, V., Mendes-Ferreira, A., Melo-Pinto, P., 2021. Application of hyperspectral imaging and deep learning for robust prediction of sugar and pH levels in wine grape berries. Sensors 21 (10), 3459. http://dx.doi.org/10.3390/s21103459.

Gouot, J.C., Smith, J.P., Holzapfel, B.P., Walker, A.R., Barril, C., 2019. Grape berry flavonoids: A review of their biochemical responses to high and extreme high temperatures. J. Exp. Bot. 70 (2), 397–423. http://dx.doi.org/10.1093/jxb/ery392.

Gutiérrez, S., Tardaguila, J., Fernández-Novales, J., Diago, M., 2019. On-the-go hyperspectral imaging for the in-field estimation of grape berry soluble solids and anthocyanin concentration. Aust. J. Grape Wine Res. 25 (1), 127–133. http://dx.doi.org/10.1111/ajgw.12376.

Hernández-Hierro, J.M., Nogales-Bueno, J., Rodríguez-Pulido, F.J., Heredia, F.J., 2013. Feasibility study on the use of near-infrared hyperspectral imaging for the screening of anthocyanins in intact grapes during ripening. J. Agric. Food Chem. 61 (41), 9804–9809. http://dx.doi.org/10.1021/jf4021637.

Iland, P., Dry, P., Proffitt, T., Tyerman, S., 2011. The Grapevine: from the Science to the Practice of Growing Vines for Wine. Patrick Iland Wine Promotions.

Isaksson, T., Næs, T., 1988. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. Appl. Spectrosc. 42 (7), 1273–1284.

Kalopesa, E., Karyotis, K., Tziolas, N., Tsakiridis, N., Samarinas, N., Zalidis, G., 2023. Estimation of sugar content in wine grapes via in situ VNIR–SWIR point spectroscopy using explainable artificial intelligence techniques. Sensors 23 (3), 1065. http://dx.doi.org/10.3390/s23031065.

Lorenz, D., Eichhorn, K., Bleiholder, H., Klose, R., Meier, U., Weber, E., 1995. Growth stages of the grapevine: Phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera)—Codes and descriptions according to the extended BBCH scale†. Aust. J. Grape Wine Res. 1 (2), 100–103. http://dx.doi.org/10.1111/j.1755-0238.1995.tb00085.x.

Lowekamp, B.C., Chen, D.T., Ibáñez, L., Blezek, D., 2013. The design of SimpleITK. Front. Neuroinformatics 7, 45. http://dx.doi.org/10.3389/fninf.2013.00045.

Lu, B., Dao, P.D., Liu, J., He, Y., Shang, J., 2020. Recent advances of hyperspectral imaging technology and applications in agriculture. Remote Sens. 12 (16), http://dx.doi.org/10.3390/rs12162659.

Matese, A., Di Gennaro, S.F., Orlandi, G., Gatti, M., Poni, S., 2022. Assessing grapevine biophysical parameters from unmanned aerial vehicles hyperspectral imagery. Front. Plant Sci. 13, 898722. http://dx.doi.org/10.3389/fpls.2022.898722.

Mattes, D., Haynor, D.R., Vesselle, H., Lewellyn, T.K., Eubank, W., 2001. Nonrigid multimodality image registration. In: Medical Imaging 2001: Image Processing. Vol. 4322, Spie, pp. 1609–1620. http://dx.doi.org/10.1117/12.431046.

Meléndez, E., Ortiz, M., Sarabia, L., Íñiguez, M., Puras, P., 2013. Modelling phenolic and technological maturities of grapes by means of the multivariate relation between organoleptic and physicochemical properties. Anal. Chim. Acta 761, 53–61. http://dx.doi.org/10.1016/j.aca.2012.11.021.

Pesarin, F., Salmaso, L., 2010. Permutation Tests for Complex Data: Theory, Applications and Software. John Wiley & Sons, http://dx.doi.org/10.1002/9780470689516.

Poni, S., Frioni, T., Gatti, M., 2023. Summer pruning in Mediterranean vineyards: is climate change affecting its perception, modalities, and effects? Front. Plant Sci. 14, http://dx.doi.org/10.3389/fpls.2023.1227628.

Poni, S., Gatti, M., Palliotti, A., Dai, Z., Duchêne, E., Truong, T.-T., Ferrara, G., Matarrese, A.M.S., Gallotta, A., Bellincontro, A., Mencarelli, F., Tombesi, S., 2018. Grapevine quality: A multiple choice issue. Sci. Hort. 234, 445–462. http://dx.doi.org/10.1016/j.scienta.2017.12.035.

Power, A., Truong, V.K., Chapman, J., Cozzolino, D., 2019. From the laboratory to the vineyard—Evolution of the measurement of grape composition using NIR spectroscopy towards high-throughput analysis. High-Throughput 8 (4), http://dx.doi.org/10.3390/ht8040021.

Prasanna, V., Prabha, T.N., Tharanathan, R.N., 2007. Fruit ripening phenomena–An overview. Crit. Rev. Food Sci. Nutr. 47 (1), 1–19. http://dx.doi.org/10.1080/10408390600976841.

Rodríguez-Pulido, F.J., Mora-Garrido, A.B., González-Miret, M.L., Heredia, F.J., 2022. Research progress in imaging technology for assessing quality in wine grapes and seeds. Foods 11 (3), 254. http://dx.doi.org/10.3390/foods11030254.

Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639. http://dx.doi.org/10.1021/ac60214a047.

Smith, G.M., Milton, E.J., 1999. The use of the empirical line method to calibrate remotely sensed data to reflectance. Int. J. Remote Sens. 20 (13), 2653–2662. http://dx.doi.org/10.1080/014311699211994.

Sousa, J.J., Toscano, P., Matese, A., Di Gennaro, S.F., Berton, A., Gatti, M., Poni, S., Pádua, L., Hruška, J., Morais, R., et al., 2022. UAV-based hyperspectral monitoring using push-broom and snapshot sensors: A multisite assessment for precision viticulture applications. Sensors 22 (17), 6574. http://dx.doi.org/10.3390/s22176574.

Tang, Y., Qiu, J., Zhang, Y., Wu, D., Cao, Y., Zhao, K., Zhu, L., 2023. Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: A review. Precis. Agric. 1–37. http://dx.doi.org/10.1007/s11119-023-10009-9.

Tsakiridis, N.L., Samarinas, N., Kokkas, S., Kalopesa, E., Tziolas, N.V., Zalidis, G.C., 2023. In situ grape ripeness estimation via hyperspectral imaging and deep autoencoders. Comput. Electron. Agric. 212, 108098. http://dx.doi.org/10.1016/j.compag.2023.108098.

Tukey, J.W., et al., 1977. Exploratory Data Analysis, vol. 2, Reading, MA.

Vrochidou, E., Bazinas, C., Manios, M., Papakostas, G.A., Pachidis, T.P., Kaburlasos, V.G., 2021. Machine vision for ripeness estimation in viticulture automation. Horticulturae 7 (9), http://dx.doi.org/10.3390/horticulturae7090282.

Xu, M., Sun, J., Cheng, J., Yao, K., Wu, X., Zhou, X., 2023. Non-destructive prediction of total soluble solids and titratable acidity in Kyoho grape using hyperspectral imaging and deep learning algorithm. Int. J. Food Sci. Technol. 58 (1), 9–21. http://dx.doi.org/10.1111/ijfs.16173.

Yaniv, Z., Lowekamp, B.C., Johnson, H.J., Beare, R., 2018. SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. J. Digit. Imaging 31 (3), 290–303. http://dx.doi.org/10.1007/s10278-017-0037-8.

Ye, W., Xu, W., Yan, T., Yan, J., Gao, P., Zhang, C., 2022. Application of near-infrared spectroscopy and hyperspectral imaging combined with machine learning algorithms for quality inspection of grape: A review. Foods 12 (1), 132. http://dx.doi.org/10.3390/foods12010132.

Zitová, B., Flusser, J., 2003. Image registration methods: a survey. Image Vis. Comput. 21 (11), 977–1000. http://dx.doi.org/10.1016/S0262-8856(03)00137-9.