Information Fusion 99 (2023) 101865

Contents lists available at ScienceDirect

Information Fusion

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

Full length article

Multi-sensor spectral fusion to model grape composition using deep learning

Salvador Gutiérrez [a,*], Juan Fernández-Novales [b,c], Teresa Garde-Cerdán [c], Sandra Marín-San Román [c], Javier Tardaguila [b,c], María P. Diago [b,c,**]

[a] Department of Computer Science and Artificial Intelligence (DECSAI), Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada (UGR), Granada, 18014, Spain
[b] University of La Rioja, Department of Agriculture and Food Science, Madre de Dios 53, Logroño, 26007, Spain
[c] Instituto de Ciencias de la Vid y del Vino (CSIC, Universidad de La Rioja, Gobierno de La Rioja), Logroño, 26006, Spain

ARTICLE INFO

Keywords:
Multi-block
Chemometrics
Spectrometer
Convolutional neural networks
Multilayer perceptrons
Spectroscopy
Amino acids
Nitrogen compounds

ABSTRACT

Spectral instruments can be useful for the rapid assessment of chemical compounds in different targets, and their use have been already reported for the modeling of grape composition comparing two spectral ranges. Still, with the increased easiness of acquiring data with several sensors, it would be valuable to explore spectral fusion techniques for the modeling with deep learning, seeking to obtain improved performance. Therefore, the objective of this work was to develop multi-sensor spectral fusion approaches for the deep learning modeling of grape composition. From 128 grape samples, two spectra per sample were acquired from two different ranges using two sensors (visible and shortwave near infrared, 570–1000 nm; and wider NIR 1100–2100 nm). From each sample, 15 grape nitrogen compounds were analyzed by wet chemistry. Three different data fusion approaches are defined using neural networks and deep learning, testing several ways of structuring and merging the input spectra. Statistical analyses supported that (i) the proposed deep learning fusion architectures performed better than single spectral range models, and (ii) neural networks have better modeling capabilities than partial least squares in spectral fusion. The results demonstrate the potential of deep learning for spectral data fusion in grape nitrogen composition regression, and potentially other traits in food and agriculture spectroscopy.

1. Introduction

The use of sensors is a critical part of chemometrics, as they provide data that can be useful for the rapid monitoring and modeling of important parameters from many samples [1]. Within the different sensing technologies, optical spectrometers have proven their effectiveness in modeling chemical compounds in many applications [2–4]. Nevertheless, it is frequent that the target parameter sought to be modeled is too complex for its effective analysis using a single instrument—in the case of spectroscopy, a single spectral range usually, as the spectral response could be present in a different range, or a combination of these. For this reason, research on the sensor or source-data fusion has been explored in chemometrics in many occasions, seeking to extract better information from the samples by the combination of different sensing techniques, spectral ranges or even appending data from chemical measurements [5]. This sensor and data fusion strategy, often referred as to multi-block analysis in chemometrics, has proven effective for many applications [6], especially when combined with classical classification and regression methods like partial least squares (PLS) [7,8].

Machine and deep learning have attracted the attention regarding the modeling of spectral data for years. While classical multivariate techniques are common in chemometrics (principal component analysis, PLS, etc.), they are limited to linear relationships between the spectra and the target chemical, although this is often sufficient. Machine learning techniques such as fully connected neural networks and support vector machines have been successfully applied for the training of prediction models from spectral data [9,10], while more recently deep learning techniques like convolutional neural networks have also worked with spectral input, testing especially one-dimensional filters that convolve over the spectral dimension [11–13]. The strengths of the backpropagation for the training of neural networks are indisputable, including the modeling of spectral data, that is featured by a general presence of collinearity between adjacent or very close wavelengths. While neural network models have been reported for spectral data, few works can be found about spectral fusion by neural networks.

* Corresponding author.
** Corresponding author at: University of La Rioja, Department of Agriculture and Food Science, Madre de Dios 53, Logroño, 26007, Spain.
E-mail addresses: salvaguti@decsai.ugr.es (S. Gutiérrez), maria-paz.diago@unirioja.es (M.P. Diago).

https://doi.org/10.1016/j.inffus.2023.101865
Received 21 April 2023; Received in revised form 26 May 2023; Accepted 29 May 2023
Available online 2 June 2023

Additionally to this, we found useful the possibility of designing and testing different deep learning architectures for spectral data fusion, under the same conditions for reliable comparison.

In food and agriculture, viticulture has benefited from the use of spectral sensors for the monitoring of plants and fruit [14]. Within the latter, the rapid assessment of chemical compounds from grape berries is a topic commonly addressed with different spectral ranges and instruments [15,16]. This grape composition assessment from spectral sensor is also very useful due to its potential of estimating several grape compounds (of different families) from the same source spectral data. Must nitrogen and amino acids composition are highly relevant in winemaking. Likewise, amino acids are key compounds for flavor metabolism, yeast growth and fermentation kinetics [17,18]. Therefore, given the already demonstrated usefulness of two separate spectral ranges for the determination of the amino acids profile in grape berries [19], the question arises as to whether spectral fusion of different spectral blocks improves modeling when compared to individual ranges.

The work presented in this paper was motivated by two main reasons: (i) the flexibility and potential of neural network architectures for modeling spectral data (including the combination of several spectral ranges in different structures), and (ii) the existence of previous models for the same purpose using single spectral blocks, separately. We hypothesize that, under the same samples and conditions, the fusion of two spectral ranges will perform better than using them separately, for the modeling of grape composition. Therefore, the objective of this paper was to develop a multi-sensor spectral fusion approach with deep learning to test if their performances are significantly improved against models trained on individual ranges. The specific objectives are:

1. To define and implement different deep learning fusion architectures.
2. To verify the performance boost of spectral fusion models vs single spectrum baseline, on the prediction of 15 grape berry nitrogen compounds.
3. To test the influence of deep learning models against partial least squares regression, both with data fusion.
4. To analyze the significance of different spectral pre-treatments, providing the best model for each grape berry compound.

2. Related works

Data fusion has been attempted in agriculture via different methodologies. One approach is the use of two or more sensors of different nature. Fusion of spectral and thermal data has been reported as a successful combination for yield prediction in spring barley [20], and also for soybean phenotyping using extreme machine learning [21]. The combination of spectroscopy and fluorescence has helped in the monitoring of mineral oil quality [22] and the classification of different tea types [23]. Finally, sensor fusion has been explored in grapevines for transpiration monitoring [24], water status assessment [25] or spatio-temporal delineation [26]. Another common approach is the use of two different – usually disjoint – spectral ranges (multi-block analysis). Crude oil classification was achieved by means of the Fourier-transform near and mid infrared spectroscopy [27], while the fusion of data from the same ranges was reported for quality traits prediction in tuber flours [28]. Ultraviolet visible and near infrared blocks was combined for the analysis of a wide range (from 800 to 2500 nm) to determine the active ingredient in deltamethrin formulation using extreme machine learning [29]. Finally, multi-block analysis has been applied using deep learning architectures for dry matter estimation in mangoes [30].

Regarding the use of deep learning to model spectral data in food and agriculture, convolutional neural networks have been used to train models for the prediction of total soluble solids in pears [31], alteration in coffee beans [32] or the determination of fungal contamination in maize [33]. In chemometrics, deep learning has been used for augmentation and classification in infrared spectroscopy [34]. A common approach has been the use of 1-dimensional convolutional operations (spectral dimension) [13,30,35,36], but newer approaches have been also attempted, like transformers [37,38] or deep generative techniques [39,40].

3. Methodology

3.1. Data collection

Grape sampling involved the harvesting of grape clusters along five different dates from a commercial vineyard in La Rioja (Spain), in the span of three months (from August to October, 2015). At each collection date, clusters were transported to laboratory in portable refrigerators and then stored at −20 °C until spectral and chemical processing. We considered the influence of freezing the samples to be of little significance based on previous studies in grapes [41,42]. Once the collection time frame finished, clusters (identified by its picking date) were thawed to select and take 35 representative berries (one sample) from top, mid and bottom parts.

Spectral signals from the samples were acquired with two spectrometers covering (i) the visible and shortwave near infrared (VIS-SWNIR range), and (ii) a wider near infrared range (WNIR range). Specifically, the VIS-SWNIR sensor was Polytec PSS 1050 working in the 570–1000 nm spectral range, while the WNIR sensor was a PSS 2120 working in the 1100–2100 nm range. Both having a spectral resolution of 2 nm, the VIS-SWNIR spectra involved 215 datapoints ($d_{\text{vis-swnir}} = 215$), and the WNIR sensor returned spectra with 501 datapoints ($d_{\text{wnir}} = 501$). Both spectrometers operate with sensor heads for reflected light capturing attached to their corresponding spectral processing unit (VIS-SWNIR or WNIR), connected by optic fiber. Each sensor head is equipped with embedded sample illumination (20 W tungsten halogen lamp), white reference W (material with a theoretical maximum light reflectance property) and dark current measurement D (accounting for the baseline signal from the device's electrical interferences, thermal fluctuations, etc.). With this, at a given wavelength λ, the reflectance is computed as:

$$R_\lambda = \frac{S_\lambda - D_\lambda}{W_\lambda - D_\lambda}, \tag{1}$$

where S_λ is the light intensity from the sample at wavelength λ. Absorbance spectra A can be computed from Eq. (1) as:

$$A_\lambda = \log(1/R_\lambda). \tag{2}$$

As each sample comprised 35 berries, and this number was too large for the sensor head to cover every berry, each sample was divided in seven batches of five berries each, taking one spectral measurement per batch. Finally, the spectrum corresponding to each sample was obtained by averaging the seven spectral measurements. A total of $N = 128$ samples were obtained.

After the acquisition of spectral data, each sample (group of 35 berries) was prepared for amino acids and total soluble solids analysis (TSS) by wet chemistry. Full details on the procedure is provided in [19,43]. A total of 22 amino acids per sample were analyzed (full description in [19]), and the same 15 compounds modeled with PLS in [19] were selected to be modeled with deep learning. Table 1 summarizes the names and statistics of each nitrogen compound modeled.

3.2. Spectral fusion proposals

Three different spectral fusion neural network architectures were designed and implemented, differing in the way both spectra are structured for input. We first describe briefly each architecture:

1. *concat*: Simple concatenation of both spectral ranges in a fully connected neural network. Section 3.2.1.

Table 1
Statistical summary of the 15 grape compounds analyzed on $N = 128$ samples and modeled. All nitrogen compounds are expressed in mg N/l, while total soluble solids is expressed in °Brix. std: Standard deviation; SEL: standard error of laboratory.

Abbreviation	Full name	min	max	mean	std	SEL
Asp	Aspartic acid	0.43	3.50	1.61	0.69	0.10
Glu	Glutamic acid	0.41	5.53	2.20	1.11	0.14
Ans	Asparagine	0.54	4.94	1.91	0.96	0.08
Gly	Glycine	0.36	2.39	1.17	0.48	0.20
Cit	Citrulline	0.00	5.38	1.11	0.92	0.15
GABA	γ-aminobutyric acid	1.45	18.23	9.13	3.08	0.51
Tyr	Tyrosine	0.12	2.29	0.83	0.50	0.01
Val	Valine	0.12	4.74	1.07	0.67	0.01
Trp	Tryptophan	0.28	7.17	3.00	1.44	0.05
Phe	Phenylalanine	0.13	4.92	1.23	0.74	0.01
Ile	Isoleucine	0.03	3.15	0.63	0.46	0.01
Leu	Leucine	0.01	6.26	1.23	0.92	0.04
Lys	Lysine	0.21	3.58	1.40	0.66	0.22
Pro	Proline	1.23	121.16	33.61	25.57	3.48
TSS	Total soluble solids	7.90	32.30	22.92	5.75	0.19

2. *parallel*: The two spectra stacked in a two-dimensional structure, using convolution operations for feature extraction and fully connected layers for regression. Section 3.2.2.
3. *split*: Each spectrum is fed to the network separately, using convolutional modules specific for each spectral range. The two separate outputs of this feature extraction is then joined and forwarded to fully connected layers for regression. Section 3.2.3.

We first present some definitions common to all architectures, and then we describe each in the following subsections. All tensors are written with capital letters, and one-dimensional tensors are structured as column vectors. With $d_{\text{vis-swnir}} = 215, d_{\text{wnir}} = 501$ (Section 3.1), the tensor $X_{\text{vis-swnir}} \in \mathbb{R}^{d_{\text{vis-swnir}}}$ is a spectrum (sample) from the visible sensor, and the tensor $X_{\text{wnir}} \in \mathbb{R}^{d_{\text{wnir}}}$ is a spectrum (sample) from the WNIR sensor, both linked with their corresponding values for the 15 grape compounds (Table 1). We define $\oplus : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n+m}$ as the operator one-dimensional tensor concatenation:

$$A \oplus B := \begin{bmatrix} x_1 & x_2 & \cdots & x_n & y_1 & y_2 & \cdots & y_m \end{bmatrix}^\top, A \in \mathbb{R}^n, B \in \mathbb{R}^m. \quad (3)$$

In all neural network function definitions, a trainable bias term is assumed to be added to each weight tensor matrix W, but it is omitted for better readability.

3.2.1. One dimensional concatenation

Tihs is a fully connected network with one hidden layer—a multilayer perceptron (Fig. 1). The input layer is the concatenation of VIS-SWNIR and WNIR spectra into a single tensor (with a size of 716 points), forwarded to a hidden layer of 1025 units (with ReLU as activation function) and finally converging into a single neuron as output.

Defined as $f_{\text{concat}} : \mathbb{R}^{d_{\text{vis-swnir}}} \times \mathbb{R}^{d_{\text{wnir}}} \to \mathbb{R}$:

$$f_{\text{concat}}\left(X_{\text{vis-swnir}}, X_{\text{wnir}}\right) = \text{relu}\left((X_{\text{vis-swnir}} \oplus X_{\text{wnir}})^\top \cdot W_h\right) \cdot W_{out}, \quad (4)$$

with $W_h \in \mathbb{R}^{d \times 1024}, d = d_{\text{vis-swnir}} + d_{\text{wnir}}$ as a hidden layer, and $W_{out} \in \mathbb{R}^{1024 \times 1}$ as the output layer, that maps the input to a single continuous value.

3.2.2. Parallel structure

A convolutional neural network with two modules: feature extraction with two-dimensional convolutional layers, and regression with fully connected layers (Fig. 2). The input spectra is rearranged into a two-dimensional parallel structure, with the VIS-SWNIR spectrum stretched and stacked over the WNIR spectrum. The convolutional and fully connected layers are partially based on network architectures previously tested on spectral data [44,45].

Defined as $f_{\text{parallel}} : \mathbb{R}^{d_{\text{vis-swnir}}} \times \mathbb{R}^{d_{\text{wnir}}} \to \mathbb{R}$:

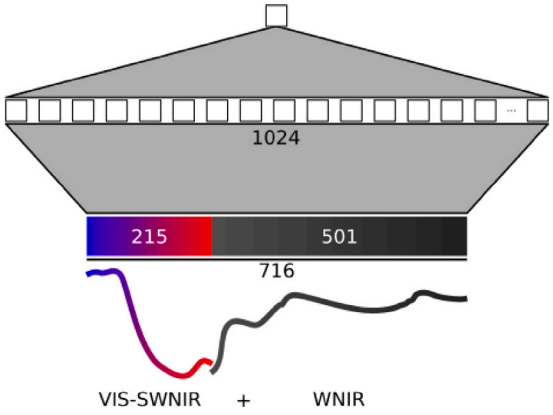

Fig. 1. Network design for the *concat* architecture. Data flow is bottom-up. The input is the concatenation of both spectral ranges, and the three layers are fully connected.

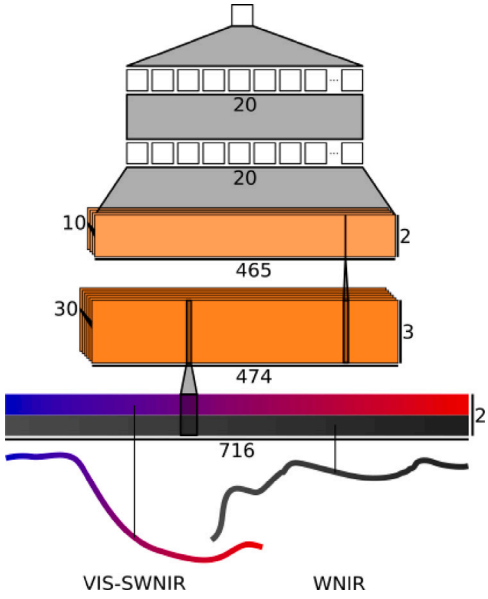

Fig. 2. Network design for the *parallel* architecture. Data flow is bottom-up. The two input spectra are reshaped into a two dimensional structure, and feed into two 2-dimensional convolutional layers and afterwards into fully connected layers.

$$f_{\text{parallel}}\left(X_{\text{vis-swnir}}, X_{\text{wnir}}\right) = \text{Dense}_{\text{p}}\left(\text{vec}\left(\text{Conv}_{\text{p}}\left(\text{stack}(X_{\text{vis-swnir}}, X_{\text{wnir}})\right)\right)\right), \quad (5)$$

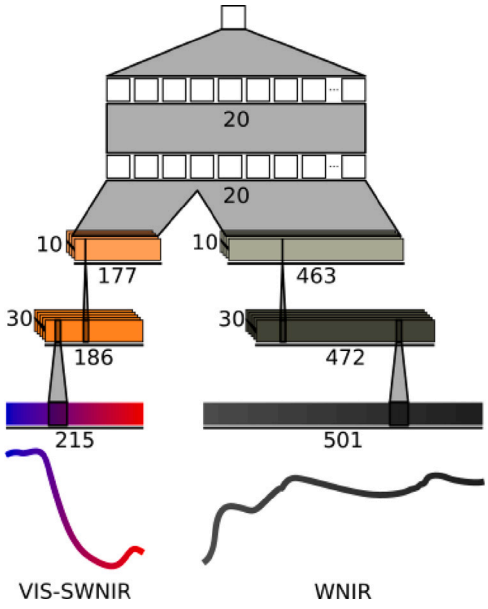

Fig. 3. Network design for the *split* architecture. Data flow is bottom-up. The two input spectra are feed into two independent groups of convolutional layers. Their respective outputs are concatenated and forwarded to fully connected layers.

where the function stack : $\mathbb{R}^{d_{\text{vis-swnir}}} \times \mathbb{R}^{d_{\text{wnir}}} \rightarrow \mathbb{R}^{2\times d_{\text{wnir}}}$ merges two separate tensors (in this case, spectrum tensors) into a two-dimensional tensor with the two spectra in parallel (upsampling $X_{\text{vis-swnir}}$ with linear interpolation to match X_{wnir} dimension d_{wnir}). The function vec vectorizes (*flattens*) the input tensor $X \in \mathbb{R}^{n\times m}$ into:

$$\text{vec}(X) := \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} & x_{2,1} & x_{2,2} & \cdots & x_{2,m} & \cdots & x_{n,m} \end{bmatrix}^\top. \tag{6}$$

The functions named Conv_p and Dense_p refer to neural networks (modules) for the feature extraction with convolution operations, and for the regression modeling with fully connected layers, respectively. Given a two spectra tensor $A \in \mathbb{R}^{2\times n}$, Conv_p is defined as:

$$\text{Conv}_\text{p}(A) = \text{relu}\big(\text{relu}\big(A \circledast W_1\big) \circledast W_2\big), \tag{7}$$

with $W_1 \in \mathbb{R}^{30\times 2\times 30}$ and $W_2 \in \mathbb{R}^{10\times 2\times 10}$ as weight parameters of two-dimensional convolution filters, with 30 and 10 one-channel filters with dimensions 2×30 and 2×10, respectively.

Given a flattened tensor $B \in \mathbb{R}^m$, Dense_p is defined as:

$$\text{Dense}_\text{p}(B) = \text{relu}\big(\text{relu}(B^\top \cdot W_1) \cdot W_2\big) \cdot W_{out}, \tag{8}$$

with $W_1 \in \mathbb{R}^{m\times 20}$, $W_2 \in \mathbb{R}^{20\times 20}$, and $W_{out} \in \mathbb{R}^{20\times 1}$ as the weights from the fully connected layers

3.2.3. Split spectral feature extraction

A convolutional neural network with two separated spectral inputs (Fig. 3), inspired by [30]. Feature extraction is done using two separated one-dimensional convolutional layers, each affecting one spectral range (VIS-SWNIR or WNIR). Regression is done with fully connected layers, after joining the two outputs from the feature extraction. As in Section 3.2.2, the convolutional and fully connected layers are partially based on networks previously tested on spectral data [44,45] (see Fig. 3).

Defined as $f_{\text{split}} : \mathbb{R}^{d_{\text{vis-swnir}}} \times \mathbb{R}^{d_{\text{wnir}}} \rightarrow \mathbb{R}$:

$$f_{\text{split}}\big(X_{\text{vis-swnir}}, X_{\text{wnir}}\big) = \text{Dense}_\text{s}\Big(\text{vec}\big(\text{Conv}_\text{s}\big(X_{\text{vis-swnir}}\big)\big) \oplus \text{vec}\big(\text{Conv}_\text{s}\big(X_{\text{wnir}}\big)\big)\Big). \tag{9}$$

The functions named Conv_s and Dense_s refer to neural networks (modules) for the feature extraction with convolution operations, and for the regression modeling with fully connected layers, respectively. Although Conv_s is used twice in Eq. (9), both are independent modules, with their own trainable parameters, so no weight sharing is done. Given a spectrum tensor $A \in \mathbb{R}^n$, and a concatenated, flattened tensor $B \in \mathbb{R}^m$, Conv_s and Dense_s are then defined as:

$$\text{Conv}_\text{s}(A) = \text{relu}\big(\text{relu}\big(A \circledast W_1\big) \circledast W_2\big), \tag{10}$$

with $W_1 \in \mathbb{R}^{30\times 30}$ and $W_2 \in \mathbb{R}^{10\times 10}$ as the weight parameters of 30 and 10 one-dimensional convolution filters of dimensions 30 and 10, respectively, and

$$\text{Dense}_\text{s}(B) = \text{relu}\big(\text{relu}(B^\top \cdot W_1) \cdot W_2\big) \cdot W_{out}, \tag{11}$$

with $W_1 \in \mathbb{R}^{m\times 20}$, $W_2 \in \mathbb{R}^{20\times 20}$, and $W_{out} \in \mathbb{R}^{20\times 1}$ as the weights from the fully connected layers.

3.3. Modeling

All neural network architectures were developed using PyTorch 1.13 on Python 3.9. The specific module architecture for each fusion approach is presented in Fig. 4. The *concat* network (Fig. 4a) involved the linear transformation of the input (the concatenation of the two spectral ranges) into 1024 neurons and, after activation with ReLU, the mapping of those into a single output scalar, with no activation. The *parallel* and *split* networks (Fig. 4b and c) forward the input into two consecutive convolutional modules (in the case of *split*, this is done separately, with independent modules), and then they flatten the data to two more linear modules (in the case of *split*, the two separated outputs from convolutions are first flattened and then merged into one structure). Each convolutional layer has attached processes for batch normalization and dropout ($p = 0.4$). All modeling layers output to a ReLU activation function, except the last neuron, that maps to a continuous scalar.

Optimization of the neural networks was done using the Adam algorithm (learning rate of 0.001), using mean squared error (MSE) with respect to the real grape composition values for the loss calculation:

$$L_{\text{MSE}} = \frac{1}{N}\sum_{k=1}^{N}(p_k - y_k)^2, \tag{12}$$

where p is the predicted value and y is the actual grape composition value. Batch size was set to 32 and the training of each model stopped at 2000 epochs. The selection of the mentioned values for hyperparameters was performed after a supervised grid search on several random data combinations (target, spectral pre-processing, etc.) with values based on our previous experience. For each model, training was done with 80% of the samples, using the remaining 20% for validation. As in [19], data partition was designed so as to samples are randomly assigned to each dataset, but ensuring the preservation of the grape compound distribution in both subsets, to assure that results from data fusion are comparable to the original publication.

Spectral data can be treated with many pre-processing techniques [46], and this can greatly influence in the performance of the trained models. For this reason, for each grape compound model, a large combination of spectral pre-processing parameters were used for the training of several models. The spectral variations explored were the following:

- **Spectral mode**: Presenting the spectrum as *reflectance* or *absorbance*.
- **Scatter correction**: Applying (*true*) standard normal variate and detrending to the spectrum, or not applying scatter correction (*false*).
- **Spectral smoothing**: Using Savitzky–Golay filtering (*true*) or not using it (*false*).

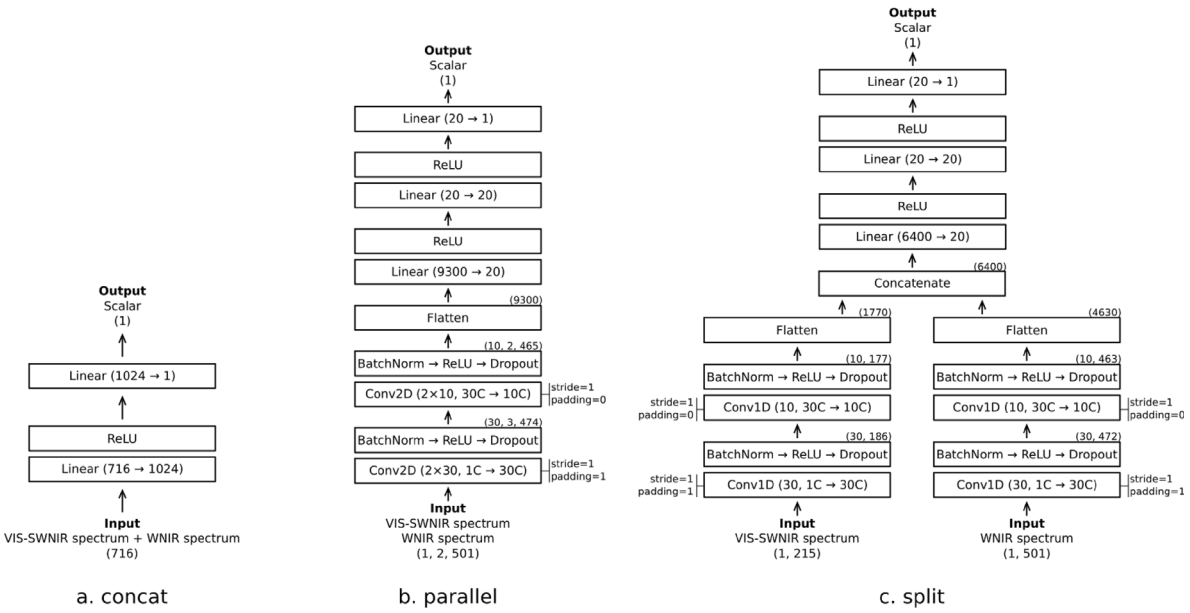

Fig. 4. Module architectures of the three fusion neural networks proposals.

Table 2
Description of the values of each parameter used for the training of the models in this work. The combination of all these parameter values for the training of models for the prediction of each one of the 15 grape composition targets makes up a total of $n = 2520$ trained models.

Parameter	Value	Description
Spectral mode	*reflectance*	The spectrum is input as reflectance (Eq. (1)).
	absorbance	The spectrum is input as absorbance (Eq. (2)).
Scatter correction	*true*	Scatter correction is applied to the spectrum.
	false	No scatter correction.
Spectral smoothing (Savitzky–Golay filtering)	*true*	The spectrum is smoothed with a Savitzky–Golay filtering.
	false	No spectral smoothing
Derivative order (only when spectral smoothing is applied)	*first*	The Savitzky–Golay filtering is done with a first-order derivative.
	second	The Savitzky–Golay filtering is done with a second-order derivative.
Window size (only when spectral smoothing is applied)	*5*	Window size of 5 for the Savitzky–Golay filtering.
	9	Window size of 9 for the Savitzky–Golay filtering.
	15	Window size of 15 for the Savitzky–Golay filtering.
Model	*concat*	Neural network architecture described in Section 3.2.1.
	parallel	Neural network architecture described in Section 3.2.2.
	split	Neural network architecture described in Section 3.2.3.
	wnir	Neural network similar to concat but only with the WNIR range.
	vis-swnir	Neural network similar to concat but only with the Vis-SWNIR range.
	pls	PLS modeling with both spectral ranges concatenated in one dimension.

- **Derivative order**: using *first* or *second* derivative order, only when Savitzky–Golay filtering is applied.
- **Window size**: smoothing window size of *5*, *9* or *15*, again only when Savitzky–Golay filtering is used.

The combination of all these spectral pre-processing parameters were used for the training of models for all the 15 grape composition targets, and for the three neural networks architectures (*concat*, *parallel* and *split*). Apart from these, additional models were trained using an architecture similar to *concat* (Section 3.2.1), but only considering VIS-SWNIR spectra as input (*vis-swnir* network), and also only considering WNIR spectra (*wnir* network). This was done to test the influence of spectral fusion vs models from individual ranges. Finally, to test the performance of neural networks fusion models against techniques used in the original work [19], PLS models (*pls*) were trained using as input the concatenation of VIS-SWNIR and WNIR spectra (like the *concat*, *vis-swnir* and *wnir* networks). The PLS implementation used was from scikit-learn 1.2.1 [47], retaining no more than seven latent variables. A summary of all the parameters used for modeling is presented in Table 2.

With all these parameters, the proper combination of them all made up a total of $n = 2520$ trained models. Training was done in parallel on two NVIDIA GeForce RTX 4090 with 24 GB of vRAM each, requiring approximately 20 h.

3.4. Statistical analyses

With all the models obtained from the combination of different architectures, pre-processing, etc. (Section 3.3), the influence of different parameters were examined using analysis of variance. For each individual model (from the $n = 2520$ total number of models), the performance results were the MSE (the loss function used in neural network optimization, Eq. (12)) and the determination coefficient (R^2), taken from the model with the lowest validation error for the 2000 epochs. Statistical tests were carried out using InfoStat software [48], version 2020, using Tukey's range test at a significance level $p = 0.05$.

Table 3
Best models per grape composition target among deep learning fusion (*concat*, *parallel* and *split*) and baseline models (models trained only with one spectral range). Baseline models refer to the best models trained with the VIS-SWNIR or WNIR range, using neural networks (fifth column group, using an architecture similar to *concat*, Section 3.3) and PLS (sixth column group), reported in [19]. The results with the highest R^2 per target are underlined. MSE: Mean squared error. SEP: Standard error of prediction, as reported in [19]. In all cases, error values are expressed in mg N/l, except for total soluble solids, expressed in °Brix.

Target	Deep learning fusion models						Baseline models			
	concat		*parallel*		*split*		Neural networks		PLS from [19]	
	MSE	R^2	MSE	R^2	MSE	R^2	MSE	R^2	SEP	R^2
Asp	0.188	0.70	0.212	0.63	0.227	0.65	0.224	0.63	0.130	0.60
Glu	0.741	0.54	0.534	0.61	0.734	0.63	1.362	0.65	0.490	0.46
Ans	0.658	0.37	0.520	0.46	0.558	0.41	0.704	0.47	0.203	0.66
Gly	0.119	0.57	0.066	0.75	0.150	0.63	0.130	0.53	0.090	0.37
Cit	0.745	0.55	0.373	0.55	1.141	0.54	0.407	0.62	0.260	0.43
GABA	6.656	0.47	5.027	0.53	5.299	0.49	6.641	0.52	6.350	0.32
Tyr	0.241	0.83	0.051	0.83	0.047	0.81	0.242	0.79	0.068	0.63
Val	0.376	0.43	0.217	0.46	0.301	0.48	0.254	0.55	0.123	0.59
Trp	0.724	0.67	0.832	0.66	0.704	0.70	0.953	0.66	0.608	0.58
Phe	0.238	0.55	0.215	0.62	0.238	0.55	0.272	0.58	0.152	0.43
Ile	0.071	0.63	0.091	0.57	0.095	0.49	0.090	0.59	0.044	0.52
Leu	0.307	0.64	0.275	0.70	0.289	0.68	0.448	0.61	0.270	0.34
Lys	0.297	0.37	0.315	0.37	0.262	0.40	0.319	0.49	0.194	0.62
Pro	248.220	0.66	156.800	0.76	251.571	0.65	279.057	0.71	306.250	0.63
TSS	3.038	0.95	2.307	0.95	2.137	0.95	2.022	0.96	2.528	0.91

4. Experimental results

All mean comparison performed and shown in this section (by analysis of variance) were partitioned and analyzed individually for each grape compound, as the interaction between nitrogen compound (Table 1) and type of model (*concat*, *parallel*, *split*, *vis-swnir*, *wnir*, *pls*) was found to be statistically significant ($p < 0.0001$).

4.1. Deep learning fusion models

The best results among fusion and baseline (individual spectral range) models are presented in Table 3, for each of the target grape nitrogen compounds. The values from the first three column groups (*concat*, *parallel* and *split*) come from models trained with both VIS-SWNIR and WNIR ranges, while the models reported in the last two column groups, "Neural networks" and "PLS", can come from either VIS-SWNIR or WNIR ranges. As the determination coefficient values can be compared among models, the results with the highest R^2 per target are underlined in Table 3.

In absolute terms, the best results came from fusion models in nine out of 15 target nitrogen compounds, with special evidence in models for predicting Asp, Gly, Tyr, Leu and Pro, that exhibited increases in R^2 with large differences in comparison to the baseline PLS models. Single range neural network models (fifth column group in Table 3) were better in three times (for Glu, Cit, TSS) and, along with the deep learning fusion models, they constituted the majority of the high-performance results (in 12 out of 15 grape composition parameters) vs PLS (the remaining three). The prediction of TSS is one case that can be considered special, as a very high performance results came from all models, demonstrating that the influence of this composition parameter is strongly present in either VIS-SWNIR and WNIR ranges, and easily modeled by any algorithm.

Given the absolute results from the three deep learning fusion proposals proposed in this work, analysis of variance was done to test if these fusion architectures–*concat*, *parallel* and *split*–did have positive influence in the performance of trained models, compared to the remaining models developed—*vis-swnir*, *wnir* and *pls*. The analyses, again done for each grape nitrogen compound, are gathered in Table 4.

In all cases, the deep learning fusion models performed better than the remaining ones (*i.e.*, the errors were lower and the R^2 values were higher). Moreover, regarding the MSE, these differences were significant in eight out of 15 targets, highlighting large differences for GABA (7.36 vs 8.63 mg N/l), Trp (1.07 vs 1.32 mg N/l) and Pro (317.68 vs 361.17 mg N/l). Regarding the determination coefficients, significant differences between means were found for all grape compounds, including those with non-significant differences in error (Glu, Tyr, Val, Ile, Lys, Pro and TSS). These results manifest that the proposed fusion architectures might benefit the prediction capabilities over models trained from individual ranges. This hypothesis was also tested previously by other authors in slightly similar ranges [30,49], reporting the benefits of data fusion vs single block modeling. The deep learning architectures were designed to handle all the data combinations available – target, pre-processing, etc.–, allowing for a versatile and adaptable approach. While it is acknowledged that the performance of these models could potentially be further enhanced through fine-tuning on individual datasets, the objective of the experiments was to ensure comparability and consistency in the evaluation process. After confirmed the superior performance from spectral fusion, fine-tuning of specific models can be a logical step for their deployment in real-time spectral applications for grape compound estimation.

4.2. Performance of deep learning against partial least squares for fusion models

After finding significant differences for many grape compounds in favor of the deep learning fusion models, these were individually compared to PLS models trained with the concatenation of both VIS-SWNIR and WNIR spectral ranges, to find if neural networks actually improve the models' prediction capability. The results of the analyses are presented in Table 4.

Regarding the MSE values, *pls* means showed larger errors in almost all cases compared to the three neural network architectures, resulting in significant differences in 10 out of 15 grape nitrogen compounds in favor of neural networks. Within these, the mean performance from models trained with *parallel* yielded, in general, lower error values than *concat* and *split* in all targets with significant differences (except for Pro and Trp), but the differences are not large, so no architecture exhibited clearly a significant increase in performance over the remaining ones.

When considering the determination coefficient (Table 5), significant differences in means were found for all compounds. Additionally, differences were in all cases balanced towards the methods (better R^2 values for *concat*, *parallel* and *split*), except for TSS, in which models trained with the *parallel* and *split* architectures achieved significantly greater outputs than the others. Apart from this, it is necessary to highlight that both *parallel* and *split* use convolutional modules (two and one-dimensional, respectively) although with different input structures.

Table 4
Analysis of variance between deep learning fusion models (*concat*, *parallel* and *split*) and the remaining ones (*vis-swnir*, *wnir* and *pls*), for error and R^2 values. DL fusion: Models trained with the proposed deep learning fusion architectures. Rest: Models trained other techniques (single spectral range neural networks, or partial least squares). Dissimilar lowercase letters indicate significant differences (ns: not significant, $p \geq 0.05$) using the Student–Neumann–Keuls test. Letters of significance are associated to increased model performance, that is descendent order of R^2 and ascendent order of mean squared error. Error values are expressed in mg N/l, except for total soluble solids, expressed in °Brix.

Target	Mean squared error			R^2		
	DL fusion	Rest	*p*-value	DL fusion	Rest	*p*-value
Asp	0.26a	0.29b	0.0001	0.54a	0.45b	<0.0001
Glu	0.93	1.05	ns	0.45a	0.30b	<0.0001
Ans	0.77a	0.85b	0.0010	0.28a	0.23b	<0.0001
Gly	0.15a	0.17b	<0.0001	0.44a	0.33b	<0.0001
Cit	0.59a	0.85b	0.0001	0.36a	0.25b	<0.0001
GABA	7.36a	8.63b	0.0002	0.36a	0.24b	<0.0001
Tyr	0.12	0.13	ns	0.70a	0.63b	<0.0001
Val	0.33	0.35	ns	0.36a	0.31b	0.0001
Trp	1.07a	1.32b	<0.0001	0.56a	0.46b	<0.0001
Phe	0.29a	0.34b	0.0005	0.45a	0.35b	<0.0001
Ile	0.18	0.18	ns	0.40a	0.34b	0.0001
Leu	0.45a	0.58b	<0.0001	0.51a	0.37b	<0.0001
Lys	0.39	0.41	ns	0.28a	0.22b	<0.0001
Pro	317.68a	361.17b	ns	0.54a	0.51b	0.0101
TSS	5.76	6.69	ns	0.89a	0.84b	0.0077

Table 5
Comparison of means between the three proposed fusion architectures and PLS fusion models. The models summarized in this table were all trained with the fusion of both spectral ranges (VIS-SWNIR and WNIR), using deep learning (*concat*, *parallel*, *split*) or partial least squares (*pls*). Dissimilar lowercase letters indicate significant differences (ns: not significant, $p \geq 0.05$) using the Student–Neumann–Keuls test. Letters of significance are associated to increased model performance, that is descendent order of R^2 and ascendent order of mean squared error. Error values are expressed in mg N/l, except for total soluble solids, expressed in °Brix.

Target	Mean squared error				*p*-value
	pls	*concat*	*parallel*	*split*	
Asp	0.28	0.27	0.27	0.25	ns
Glu	1.17	0.86	0.91	1.03	ns
Ans	0.90a	0.81b	0.73c	0.77bc	<0.001
Gly	0.17a	0.15b	0.14b	0.14b	<0.001
Cit	1.38a	0.71b	0.52b	0.54b	<0.001
GABA	9.76a	8.37b	6.80c	6.91c	<0.001
Tyr	0.11ab	0.15a	0.10b	0.12ab	0.032
Val	0.35	0.33	0.34	0.32	ns
Trp	1.52a	1.18b	1.05bc	0.97c	<0.001
Phe	0.41a	0.29b	0.26b	0.32b	<0.001
Ile	0.21	0.24	0.16	0.13	ns
Leu	0.69a	0.51b	0.42b	0.42b	<0.001
Lys	0.47a	0.43a	0.37b	0.38b	<0.001
Pro	367.1a	336.6b	315.2bc	301.2c	<0.001
TSS	4.69	6.78	4.36	6.14	ns
Target	R^2				*p*-value
	pls	*concat*	*parallel*	*split*	
Asp	0.46b	0.54a	0.53a	0.55a	<0.001
Glu	0.20c	0.41b	0.45a	0.48a	<0.001
Ans	0.19b	0.27a	0.29a	0.27a	<0.001
Gly	0.31b	0.40a	0.46a	0.45a	<0.001
Cit	0.10c	0.31b	0.38a	0.38a	<0.001
GABA	0.16c	0.30b	0.38a	0.40a	<0.001
Tyr	0.57b	0.69a	0.71a	0.71a	<0.001
Val	0.26c	0.33b	0.37a	0.38a	<0.001
Trp	0.35c	0.52b	0.57a	0.60a	<0.001
Phe	0.22c	0.42b	0.49a	0.44b	<0.001
Ile	0.25c	0.36b	0.42a	0.42a	<0.001
Leu	0.26b	0.47a	0.53a	0.52a	<0.001
Lys	0.14c	0.26b	0.29a	0.29a	<0.001
Pro	0.46b	0.52a	0.54a	0.55a	<0.001
TSS	0.87bc	0.84c	0.92a	0.91ab	<0.001

The results from Table 5 suggest that the input structure of the VIS-SWNIR and WNIR spectra is not relevant in this case, as no significant differences were found between the two convolutional architectures. The only exception to this was Phe, that benefited significantly more from *parallel* ($R^2 = 0.49$) than *split* ($R^2 = 0.44$).

These results show that, at the same input (the fusion of both spectral ranges), artificial neural networks performed better than PLS in both error and R^2 values, suggesting that neural networks are able to extract more information from spectra than linear methods like PLS. Previous work can be found also reporting better performance from neural networks vs PLS, in spectral regression modeling [13] and classification [50,51], also including one-dimensional convolutional neural networks [11,12]. This reinforces the hypothesis that non-linear methods perform better than PLS. Still, this behavior was not found for all grape nitrogen compound models, and PLS, able to find linear relations between the spectrum and the target, is powerful enough to model certain chemical components.

4.3. Influence of spectral pre-processing in fusion models for each grape nitrogen compound

Finally, the influence of the different spectral pre-processing techniques tested were reported using analysis of variance. The results are found in Table 6 (for better readability, only MSE values were reported, and *p*-values were not provided). This table gathers the means of all the trained models, separated and statistically analyzed by the different spectral pre-processing values presented in Table 2. The influence of using reflectance vs absorbance was not significant for any compound (data not shown). This can be considered an expected outcome as, even when the relation between reflection and absorbance is not linear (Eq. (1)), modeling techniques are still able to identify key features related to each target. This is more evident when using neural networks, as they model non-linear relationships due to the use of activation functions.

Applying scatter was mainly not relevant, and no significant differences between means were found in 12 out of 15 grape compounds. Additionally, in the remaining three targets, no clear trend can be derived (using or omitting scatter correction), so it can be concluded that it has little to none influence in the results, for these ranges and spectral acquisition conditions. The use of Savitzky–Golay filtering was marginally more meaningful, as it provided better results (lower MSE values in Table 6) depending on its use, and the derivative order applied. In general, better results came from first derivative, but this cannot be taken lightly as a general rule. Finally, regarding the window size, the general trend is balanced towards the extreme values, as the best performance per target is often found when using 5 or 15 as window size.

Table 6
Comparison of means of the mean squared error metrics attending to the different spectral pre-processing techniques tested. Dissimilar lowercase letters indicate significant differences using the Student–Neumann–Keuls test ($p \geq 0.05$). Letters of significance are associated to increased model performance, that is descendent order of R^2 and ascendent order of error. Values are expressed in mg N/l, except for total soluble solids, expressed in °Brix.

Target	Scatter correction		Savitzky–Golay			Window size		
	false	*true*	*false*	*first*	*second*	*5*	*9*	*15*
Asp	0.28	0.27	0.32b	0.26a	0.28a	0.29b	0.27a	0.27a
Glu	0.99	1.00	0.96ab	0.87a	1.12b	1.05	0.98	0.91
Ans	0.79	0.82	0.89b	0.77a	0.82ab	0.84	0.78	0.78
Gly	0.16	0.16	0.18b	0.15a	0.16a	0.16	0.16	0.15
Cit	0.69	0.75	0.7	0.78	0.67	0.71	0.72	0.73
GABA	7.84	8.15	8.16	8.19	7.75	7.56a	7.98ab	8.67b
Tyr	0.11b	0.15a	0.13	0.13	0.13	0.11	0.14	0.14
Val	0.35	0.33	0.33	0.32	0.37	0.33	0.32	0.38
Trp	1.22	1.17	1.41b	1.22a	1.10a	1.18ab	1.10a	1.32b
Phe	0.32	0.31	0.29	0.33	0.31	0.3	0.3	0.34
Ile	0.18	0.18	0.16	0.18	0.18	0.15	0.22	0.19
Leu	0.55a	0.48b	0.67b	0.49a	0.49a	0.54	0.47	0.52
Lys	0.4	0.41	0.4	0.4	0.41	0.41	0.39	0.4
Pro	354.63a	324.22b	316.14	343.31	343.31	327.36a	329.07a	367.88b
TSS	5.84	6.6	5.39a	3.98a	8.74b	8.06b	5.31a	4.38a

5. Conclusion

This work proposed three deep neural network architectures for the spectral data fusion modeling of grape nitrogen composition, aiming to demonstrate that the performance metrics improved models trained from individual spectral ranges, previously tested. The results obtained support this hypothesis, and evidence that deep learning is able to benefit from data fusion for grape composition regression, and potentially for other traits in food and agriculture.

Related to the specific objectives defined in Section 1, it can be concluded that:

1. Deep neural networks favor the design of different architectures for spectral data fusion, varying the ways spectral blocks are structured.
2. Deep learning spectral fusion models performed significantly better than models trained from single spectrum blocks, confirming the improvement of spectral fusion for the grape composition modeling, and opening the fusion proposals for other targets.
3. Neural networks exhibited significant improvement when compared to partial least squares, when using spectral fusion, in most of the grape target compounds.
4. Spectral pre-processing had little influence in the outcomes, as they did not provide significantly large differences in performance metrics.

CRediT authorship contribution statement

Salvador Gutiérrez: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Juan Fernández-Novales:** Data curation, Writing – review & editing, Supervision. **Teresa Garde-Cerdán:** Investigation. **Sandra Marín-San Román:** Investigation. **Javier Tardaguila:** Resources. **María P. Diago:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Salvador Gutiérrez reports financial support was provided by Spanish State Research Agency.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work has been supported by the Spanish State Research Agency through project PID2019-105381GA-I00 (iScience). Funding for open access charge: Universidad de Granada / CBUA.

References

[1] L.L. Simon, H. Pataki, G. Marosi, F. Meemken, K. Hungerbuhler, A. Baiker, S. Tummala, B. Glennon, M. Kuentz, G. Steele, et al., Assessment of recent process analytical technology (PAT) trends: a multiauthor review, Org. Process Res. Dev. 19 (1) (2015) 3–62.

[2] A. Rohman, A. Windarsih, E. Lukitaningsih, M. Rafi, K. Betania, N.A. Fadzillah, The use of FTIR and Raman spectroscopy in combination with chemometrics for analysis of biomolecules in biomedical fluids: A review, Biomed. Spectrosc. Imaging 8 (3–4) (2019) 55–71.

[3] Y. Xu, P. Zhong, A. Jiang, X. Shen, X. Li, Z. Xu, Y. Shen, Y. Sun, H. Lei, Raman spectroscopy coupled with chemometrics for food authentication: A review, TRAC Trends Anal. Chem. 131 (2020) 116017.

[4] E. Boichenko, D. Kirsanov, Optical spectroscopy and chemometrics in intraoperative tumor margin assessment, TRAC Trends Anal. Chem. (2023) 116955.

[5] I. Måge, E. Menichelli, T. Næs, Preference mapping by PO-PLS: separating common and unique information in several data blocks, Food Quality Pref. 24 (1) (2012) 8–16.

[6] 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.

[7] P. Mishra, J.M. Roger, D.N. Rutledge, A. Biancolillo, F. Marini, A. Nordon, D. Jouan-Rimbaud-Bouveresse, MBA-GUI: A chemometric graphical user interface for multi-block data visualisation, regression, classification, variable selection and automated pre-processing, Chemometr. Intell. Lab. Syst. 205 (2020) 104139.

[8] K. Haroon, A. Arafeh, S. Cunliffe, P. Martin, T. Rodgers, Ć. Mendoza, M. Baker, Comparison of individual and integrated inline raman, near-infrared, and mid-infrared spectroscopic models to predict the viscosity of micellar liquids, Appl. Spect. 74 (7) (2020) 819–831.

[9] S. Liu, F. Dong, J. Hao, L. Qiao, J. Guo, S. Wang, R. Luo, Y. Lv, J. Cui, Combination of hyperspectral imaging and entropy weight method for the comprehensive assessment of antioxidant enzyme activity in tan mutton, Spectrochim. Acta A (2023) 122342.

[10] J. Li, L. Zhang, F. Zhu, Y. Song, K. Yu, Y. Zhao, Rapid qualitative detection of titanium dioxide adulteration in persimmon icing using portable raman spectrometer and machine learning, Spectrochim. Acta A 290 (2023) 122221.

[11] X. Chen, G. Cheng, S. Liu, S. Meng, Y. Jiao, W. Zhang, J. Liang, W. Zhang, B. Wang, X. Xu, et al., Probing 1D convolutional neural network adapted to near-infrared spectroscopy for efficient classification of mixed fish, Spectrochim. Acta A 279 (2022) 121350.

[12] K. Kawamura, T. Nishigaki, A. Andriamananjara, H. Rakotonindrina, Y. Tsujimoto, N. Moritsuka, M. Rabenarivo, T. Razafimbelo, Using a one-dimensional convolutional neural network on visible and near-infrared spectroscopy to improve soil phosphorus prediction in Madagascar, Remote Sens. 13 (8) (2021) 1519.

[13] K.A. Einarson, A. Baum, T.B. Olsen, J. Larsen, I. Armagan, P.A. Santacoloma, L.K. Clemmensen, Predicting pectin performance strength using near-infrared spectroscopic data: A comparative evaluation of 1-D convolutional neural network, partial least squares, and ridge regression modeling, J. Chemometr. 36 (2) (2022) e3348.

[14] E. van Wyngaard, E. Blancquaert, H. Nieuwoudt, J.L. Aleixandre-Tudo, Infrared spectroscopy and chemometric applications for the qualitative and quantitative investigation of grapevine organs, Front. Plant Sci. 12 (2021) 723247.

[15] J. Porep, A. Mattes, M. Pour Nikfardjam, D.R. Kammerer, R. Carle, Implementation of an on-line near infrared/visible (nir/vis) spectrometer for rapid quality assessment of grapes upon receival at wineries, Aust. J. Grape Wine Res. 21 (1) (2015) 69–79.

[16] A. Power, V.K. Truong, J. Chapman, D. Cozzolino, From the laboratory to the vineyard—evolution of the measurement of grape composition using nir spectroscopy towards high-throughput analysis, High-Throughput 8 (4) (2019) 21.

[17] S.-J. Bell, P.A. Henschke, Implications of nitrogen nutrition for grapes, fermentation and wine, Aust. J. Grape Wine Res. 11 (3) (2005) 242–295.

[18] T. Garde-Cerdán, C. Ancín-Azpilicueta, Effect of the addition of different quantities of amino acids to nitrogen-deficient must on the formation of esters, alcohols, and acids during wine alcoholic fermentation, LWT-Food Sci. Technol. 41 (3) (2008) 501–510.

[19] J. Fernández-Novales, T. Garde-Cerdán, J. Tardáguila, G. Gutiérrez-Gamboa, E.P. Pérez-Álvarez, M.P. Diago, Assessment of amino acids and total soluble solids in intact grape berries using contactless Vis and NIR spectroscopy during ripening, Talanta 199 (2019) 244–253.

[20] P. Rischbeck, S. Elsayed, B. Mistele, G. Barmeier, K. Heil, U. Schmidhalter, Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley, Eur. J. Agron. 78 (2016) 44–59.

[21] M. Maimaitijiang, A. Ghulam, P. Sidike, S. Hartling, M. Maimaitiyiming, K. Peterson, E. Shavers, J. Fishman, J. Peterson, S. Kadam, et al., Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine, ISPRS J. Photogramm. Remote Sens. 134 (2017) 43–58.

[22] M.S. Godinho, M.R. Blanco, F.F.G. Neto, L.M. Lião, M.M. Sena, R. Tauler, A.E. de Oliveira, Evaluation of transformer insulating oil quality using NIR, fluorescence, and NMR spectroscopic data fusion, Talanta 129 (2014) 143–149.

[23] A. Dankowska, W. Kowalewski, Tea types classification with data fusion of UV–Vis, synchronous fluorescence and NIR spectroscopies and chemometric analysis, Spectrochim. Acta A 211 (2019) 195–202.

[24] K.A. Semmens, M.C. Anderson, W.P. Kustas, F. Gao, J.G. Alfieri, L. McKee, J.H. Prueger, C.R. Hain, C. Cammalleri, Y. Yang, et al., Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach, Remote Sens. Environ. 185 (2016) 155–170.

[25] N. Ohana-Levi, I. Zachs, N. Hagag, L. Shemesh, Y. Netzer, Grapevine stem water potential estimation based on sensor fusion, Comput. Electron. Agric. 198 (2022) 107016.

[26] E. Anastasiou, A. Castrignanò, K. Arvanitis, S. Fountas, A multi-source data fusion approach to assess spatial–temporal variability and delineate homogeneous zones: A use case in a table grape vineyard in greece, Sci. Total Environ. 684 (2019) 155–163.

[27] M.K. Moro, E.V. de Castro, W. Romão, P.R. Filgueiras, Data fusion applied in near and mid infrared spectroscopy for crude oil classification, Fuel 340 (2023) 127580.

[28] L.M. Kandpal, A.M. Mouazen, R.E. Masithoh, P. Mishra, S. Lohumi, B.-K. Cho, H. Lee, Sequential data-fusion of near-infrared and mid-infrared spectroscopy data for improved prediction of quality traits in tuber flours, Infrared Phys. Technol. 127 (2022) 104371.

[29] Q. Li, Y. Huang, J. Zhang, S. Min, A fast determination of insecticide deltamethrin by spectral data fusion of UV–vis and NIR based on extreme learning machine, Spectrochim. Acta A 247 (2021) 119119.

[30] P. Mishra, D. Passos, Deep multiblock predictive modelling using parallel input convolutional neural networks, Anal. Chim. Acta 1163 (2021) 338520.

[31] X. Yu, H. Lu, D. Wu, Development of deep learning method for predicting firmness and soluble solid content of postharvest korla fragrant pear using vis/nir hyperspectral reflectance imaging, Postharvest Biol. Technol. 141 (2018) 39–49.

[32] S.S.N. Chakravartula, R. Moscetti, G. Bedini, M. Nardella, R. Massantini, Use of convolutional neural network (cnn) combined with ft-nir spectroscopy to predict food adulteration: A case study on coffee, Food Control 135 (2022) 108816.

[33] B. Wang, J. Deng, H. Jiang, Markov transition field combined with convolutional neural network improved the predictive performance of near-infrared spectroscopy models for determination of aflatoxin b1 in maize, Foods 11 (15) (2022) 2210.

[34] U. Blazhko, V. Shapaval, V. Kovalev, A. Kohler, Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra, Chemometr. Intell. Lab. Syst. 215 (2021) 104367.

[35] X. Zhang, T. Lin, J. Xu, X. Luo, Y. Ying, Deepspectra: An end-to-end deep learning approach for quantitative spectral analysis, Anal. Chim. Acta 1058 (2019) 48–57.

[36] E.R.K. Neo, J.S.C. Low, V. Goodship, K. Debattista, Deep learning for chemometric analysis of plastic spectral data from infrared and raman databases, Resour. Conserv. Recy. 188 (2023) 106718.

[37] P. Fu, Y. Wen, Y. Zhang, L. Li, Y. Feng, L. Yin, H. Yang, Spectratr: A novel deep learning model for qualitative analysis of drug spectroscopy based on transformer structure, J. Innov. Opt. Health Sci. 15 (03) (2022) 2250021.

[38] B. Yun, Y. Wang, J. Chen, H. Wang, W. Shen, Q. Li, Spectr: Spectral transformer for hyperspectral pathology image segmentation, 2021, arXiv preprint arXiv: 2103.03604.

[39] Y. Kim, W. Lee, Distributed raman spectrum data augmentation system using federated learning with deep generative models, Sensors 22 (24) (2022) 9900.

[40] Z. Cao, S. Zhang, Y. Liu, C.J. Smith, A.M. Sherman, Y. Hwang, G.J. Simpson, Spectral classification by generative adversarial linear discriminant analysis, Anal. Chim. Acta (2023) 341129.

[41] D. Cozzolino, W. Cynkar, R. Dambergs, L. Janik, M. Gishen, Effect of both homogenisation and storage on the spectra of red grapes and on the measurement of total anthocyanins, total soluble solids and ph by visual near infrared spectroscopy, J. Near Infrared Spectrosc. 13 (4) (2005) 213–223.

[42] W. Cynkar, D. Cozzolino, R.G. Dambergs, The effect of sample storage and homogenisation techniques on the chemical composition and near infrared spectra of white grapes, Food Res. Int. 42 (5–6) (2009) 653–658.

[43] T. Garde-Cerdán, R. López, J. Portu, L. González-Arenzana, I. López-Alfaro, P. Santamaría, Study of the effects of proline, phenylalanine, and urea foliar application to tempranillo vineyards on grape amino acid content. Comparison with commercial nitrogen fertilisers, Food Chem. 163 (2014) 136–141.

[44] S. Gutiérrez, A. Wendel, J. Underwood, Ground based hyperspectral imaging for extensive mango yield estimation, Comput. Electron. Agric. 157 (2019) 126–135.

[45] S. Gutiérrez, A. Wendel, J. Underwood, Spectral filter design based on in-field hyperspectral imaging and machine learning for mango ripeness estimation, Comput. Electron. Agric. 164 (2019) 104890.

[46] Å. Rinnan, F. Van Den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TRAC Trends Anal. Chem. 28 (10) (2009) 1201–1222.

[47] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[48] J. Di Rienzo, M. Balzarini, L. Gonzalez, F. Casanoves, M. Tablada, C. Walter Robledo, Infostat: software para análisis estadístico, Universidad Nacional de Córdoba, Argentina, 2010.

[49] Z. Liu, S. Yang, Y. Wang, J. Zhang, Multi-platform integration based on NIR and UV–Vis spectroscopies for the geographical traceability of the fruits of Amomum tsao-ko, Spectrochim. Acta A 258 (2021) 119872.

[50] S. Gutierrez, J. Tardaguila, J. Fernandez-Novales, M.P. Diago, Support vector machine and artificial neural network models for the classification of grapevine varieties using a portable NIR spectrophotometer, PLoS One 10 (11) (2015) e0143197.

[51] W. Liu, J. Liu, J. Jiang, Y. Li, Comparison of partial least squares-discriminant analysis, support vector machines and deep neural networks for spectrometric classification of seed vigour in a broad range of tree species, J. Near Infrared Spectrosc. 29 (1) (2021) 33–41.