One-dimensional convolutional neural networks design for fluorescence spectroscopy with prior knowledge: explainability techniques applied to olive oil fluorescence spectra

One-dimensional convolutional neural networks design for fluorescence spectroscopy with prior knowledge: explainability techniques applied to olive oil fluorescence spectra

Francesca Venturini[a,b], Umberto Michelucci[b], Michela Sperti[c], Arnaud Gucciardi[b,d], and Marco A. Deriu[c]

[a]Institute of Applied Mathematics and Physics, Zurich University of Applied Sciences, Technikumstrasse 9, 8401 Winterthur, Switzerland
[b]TOELT llc, Machine Learning Research and Development, Birchlenstr. 25, 8600 Dübendorf, Switzerland
[c]PolitoBIOMed Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy
[d]Artificial Intelligence Laboratory, University of Ljubljana, Ljubljana, Slovenia

ABSTRACT

Optical spectra, and particularly fluorescence spectra, contain a large quantity of information about the substances and their interaction with the environment. It is of great interest, therefore, to try to extract as much of this information as possible, as optical measurements can be easy, non-invasive, and can happen in-situ making the data collection a very appealing method of gathering knowledge. Artificial neural networks are known for their feature extraction capabilities and are therefore well suited for this challenge. In this work, inspired by convolutional neural network (CNN) architectures in 2D and their success with images, a novel approach using one-dimensional convolutional neural networks (1D-CNN) is used to extract information on the measured spectra by using explainability techniques. The 1D-CNN architecture has as input the entire fluorescence spectrum and takes advantage in its design of prior knowledge about the instrumentation and sample characteristics as, for example, spectrometer resolution or the expected number of relevant features in the spectrum. Even if network performance is good, it remains an open question if the features used for the predictions make sense from a physical and chemical point of view and if they match what is known from existing studies. This work studies the output of the convolutional layers, known as feature maps, to understand which features the network has effectively used for the predictions, and thus which part of the measured spectra contains the relevant information about the phenomena at the basis of what has to be predicted. The proposed approach is demonstrated by applying it to the determination of the UV absorbance at 232 nm, K_{232}, from fluorescence spectra using a dataset of 18 Spanish olive oils, which were chemically analysed from certified laboratories. The 1D-CNN successfully predicts the parameter K_{232} and enables, by studying feature maps, the clear identification of the relevant spectral features. The main contributions of this work are two. Firstly, it describes how designing the neural network architecture with prior knowledge (spectrometer resolution, etc.) will help the network in learning features that have a clear connection to the chemical composition of the substances, and thus are clearly explainable. Secondly, it shows how, in the case of olive oil, the identified features match perfectly the relevant features known from existing previous studies, thus confirming that the network is learning from the underlying chemical process.

Keywords: Fluorescence spectroscopy; fluorescence sensor; olive oil; machine learning; artificial neural networks; quality control; explainability; convolutional neural networks

Send correspondence to F. Venturini: francesca.venturini@zhaw.ch

1. INTRODUCTION

The use of artificial neural networks has shown a huge potential in almost all fields. In particular, its potential has been studied in various settings in spectroscopy, as for example in infrared and Raman spectroscopy.[1–3] A particular advantage of artificial neural networks is that there is no need of pre-processing the data manually to extract relevant features, given their feature-extraction capacity.[4,5] But this advantage, at the same time, has also been criticised for the models' black-box nature.[6] Even if models learn to predict specific quantities with high performance, it typically remains an unexplored question what information of the input has been used, and if (and how) it is possible to extract this information from the models themselves. Additionally, knowing which features are relevant may shed new light on the underlying physical or chemical processes that are linked with the parameters to be predicted. When dealing with spectroscopic data, an additional difficulty is that the inputs used by the neural networks are spectra, in other words. long arrays (for example in this work 1024 values) of real numbers, each representing the light intensity at a given wavelength. Thus, typical feature importance analysis methods, as forward or backward selection,[7] are not suited for this kind of data. Various different approaches have been studied, as interval partial least squares,[8] genetic algorithms[9] and others. Unfortunately, all have been criticised when dealing with CNN with highly-dimensional inputs for various downsides[10,11] as low reproducibility, risk of overfitting, and the loss of important features. Note that in spectra, features are most likely to be wavelength ranges and not single pixels. Thus, feature selection algorithms should identify groups of pixels, or in other words, wavelength ranges in the spectrum.

An artificial neural network is normally made by a sequence of layers, with the idea being that the first layer extracts rough data features, and the subsequent layers learn more high-level abstractions. The network used in this work is a one-dimensional convolutional neural network (1D-CNN) (described in detail in Section 2.3). Therefore, in this case, it is the first convolutional layer that extracts relevant data patterns from raw spectra.[2,12] Thus, by studying and visualising the outputs of the first convolution layer (called feature maps), light can be shed on which spectroscopic features (identified by wavelength ranges) are identified by the network as relevant and useful for the final prediction. In particular, it is instructive to compare feature maps with the raw spectra, since this will clearly identify wavelength windows considered relevant by the trained neural network.

The contribution of this work are two. Firstly, a 1D-CNN designed with prior information on the experimental setup is trained to extract the absorption coefficient K_{232} from olive oil samples and the feature maps of the first convolutional layer are visualised and compared to the raw input spectra. Secondly, the wavelength ranges identified as relevant by the trained models are identified, discussed and compared with what is expected from previous studies.

2. MATERIAL AND METHODS

2.1 Olive Oils

In this work, a total of 22 olive oil samples were analyzed. They were obtained from the 2019-2020 harvest and provided by the producer Conde de Benalúa, Granada, Spain. Oils quality assessment was performed by the producer according to the current European regulation for the commercial classification into EVOO, virgin olive oil (VOO), and lampante olive oil (LOO) categories.[13] The list of the samples is reported in a previous work[14] with the values of the chemical parameters (acidity, peroxide value, K_{270}, K_{232}, and ethyl esters) that have been obtained by chemical analysis of certified laboratories.

The fluorescence spectrum of each olive oil was acquired using a miniaturized low-cost sensor, which was described in detail in a previous work[15] under an excitation at 397 nm. Each of the 22 samples was measured 20 times. All the measurements in this work were done on undiluted samples. An overview of the measured spectra can be found in Figure 2.

2.2 Dataset Preparation

The 22 samples were measured 20 times thus resulting in a total of 440 spectra. The final dataset, therefore, is made of 440 arrays, each having 1024 values, each corresponding to a pixel of the detector inside the sensor. The signal at each pixel is the measured fluoresce intensity after the background subtraction, and normalisation to have an average of zero and a standard deviation of one were applied. This normalisation was applied to make the input data small enough to avoid numerical problems during the training phase of the neural networks.[16]

2.3 1D-Convolutional Neural Network Model

The used 1D-CNN[17–19] network is made of one convolutional layer (6 filters of size 40), one max pooling layer (size 8), one convolutional layer (4 filters of size 20), one dropout layer (rate 0.5), two subsequent flatten layers (4 neurons) and a final neuron for regression. Apart from the output neuron that uses the identity activation functions, all other layers use the ReLU (Rectified Linear Unit) activation function. The network architecture can be found in Figure 1. Additionally, the architecture exploits physics-informed knowledge on the regression task. More specifically, since the spectrometer resolution is of the order of 30 pixels, 40 pixel-wide filters in the first convolutional layer were chosen, due to the fact that changes in the signal in intervals less than 30 pixel-wide are probably highly affected by (and convoluted with) the spectrometer response function.

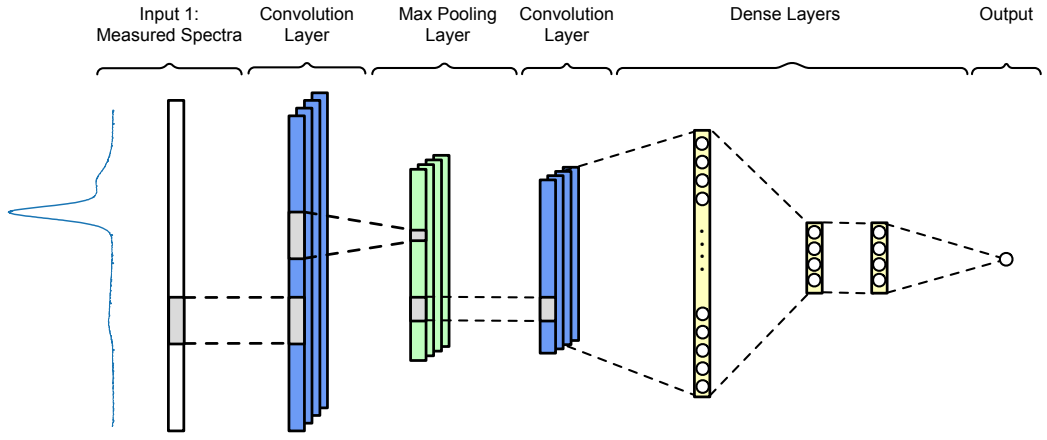

Figure 1. The neural network architecture used in this work. It is made of one convolutional layer (6 filters of size 40), one max pooling layer (size 8), one convolutional layer (4 filters of size 20), one dropout layer (rate 0.5), two subsequent flatten layers (4 neurons) and a final neuron for regression. Apart from the output neuron that uses the identity activation functions, all other layers use the ReLU (Rectified Linear Unit) activation function.

The loss function used for the training was the Mean Squared Error (MSE),[16] as reported in Equation 1, where $y^{[i]}_{\text{predicted}}$ is the predicted and $y^{[i]}_{\text{measured}}$ is the measured parameter value of the i^{th} input respectively.

$$\text{MSE}(X) = \frac{1}{N}\sum_{i=0}^{N}(y^{[i]}_{\text{predicted}} - y^{[i]}_{\text{measured}})^2 \tag{1}$$

N indicates the size of the dataset on which the metric is being evaluated. The 1D-CNN was trained using the optimiser Adam[20] and the code was implemented using the TensorFlow ^{TM} Python library. All the models were trained with backpropagation. The network has been trained for 10000 epochs with a mini-batch size of 64. The results shown are for the model that has the lowest MSE on the training dataset.

Due to the small dataset size, a leave-one-out cross-validation approach was used as described in great detail in a larger previous study.[14] Therefore, the 1D-CNN was trained on all the oils except one and this was repeated for all the oils. The results presented in this work focus on the feature maps of the convolutional layers, when the 1D-CNN was trained to predict the parameter K_{232}, and either the oil labeled D03 (quality VOO) or the oil D81 (quality EVOO) were used for the validation.

3. RESULTS AND DISCUSSION

The raw fluorescence spectra of selected EVOOs, VOOs, and LOOs are shown in Fig. 2. The spectra in all panels are just one single spectrum with the dark background subtracted, without averaging or smoothing. In the figure, the dashed line at 397 nm marks the excitation light of the LED, which is still detected by the

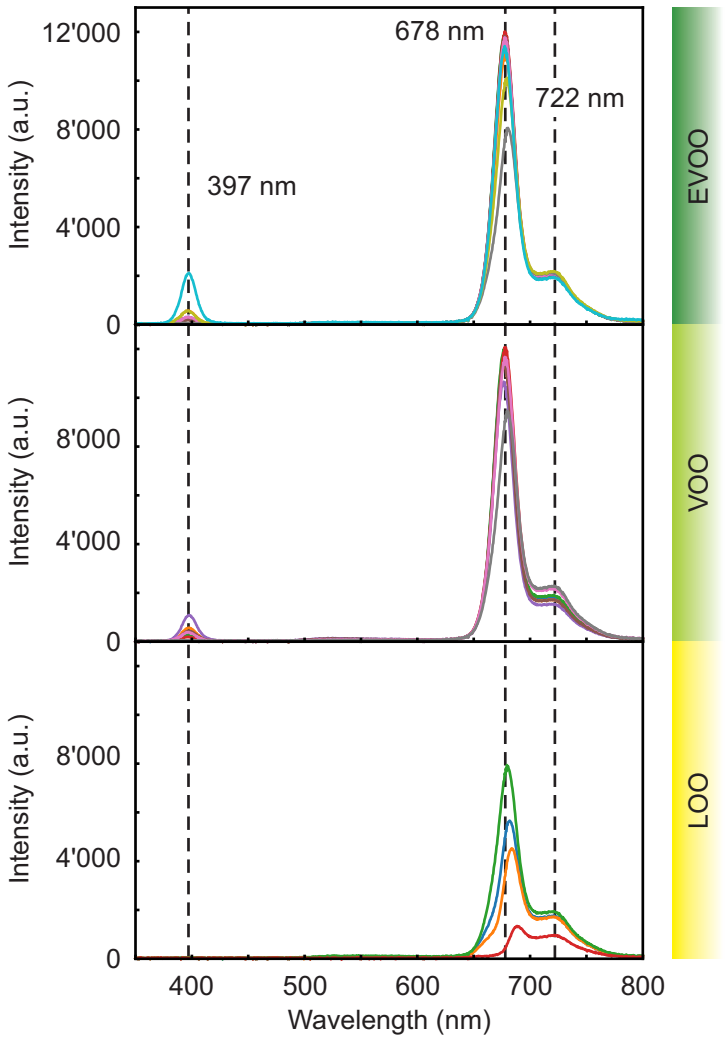

Figure 2. Fluorescence emission spectra of the measured olive oils excited at 397 nm divided in the quality classes EVOO, VOO and LOO. Each curve shows a single spectrum without averaging or smoothing.

spectrometer, despite the 90° geometry, due to the scattering on the glass vial and inside the oil. This stray light from the sample is strongest in the EVOO samples, weaker in the VOO ones, and absent in all the LOO samples, indicating that a stronger absorption in the UV is present in LOO, as expected, probably arising from oxidation products present in the oils.

In Figures 3 and 4 the feature maps of the first convolutional layers are reported for the oil D81 and D03 respectively. All the feature maps (black line) were normalised by dividing them by the maximum of the feature map 3 (the one with the highest values). After that, all were scaled so that Feature Map 3 has the same maximum as the measured spectra for readability. This was done to maintain the relative scale of the feature maps. The red line is the input measured spectrum (of the oil left out, namely D81 and D03). The black line is the feature map. The violet regions highlight the pixels regions where the feature map values are higher than 5% of their maximum. In other words, the violet region highlights the spectral regions that the 1D-CNN considers relevant (and uses) for the prediction. Note that for the two oils the order of the feature maps is not the same. This is due to the fact that the 1D-CNN predictions are invariant with respect to the exchange of filters. Thus, feature maps with the same meaning (for example Feature Map 2 for the oil D81 and Feature Map 4 for the oil D03) may appear in a different position in the feature map list.

Let us now discuss the identified relevant spectral regions. To do this let us refer to the feature maps of oil

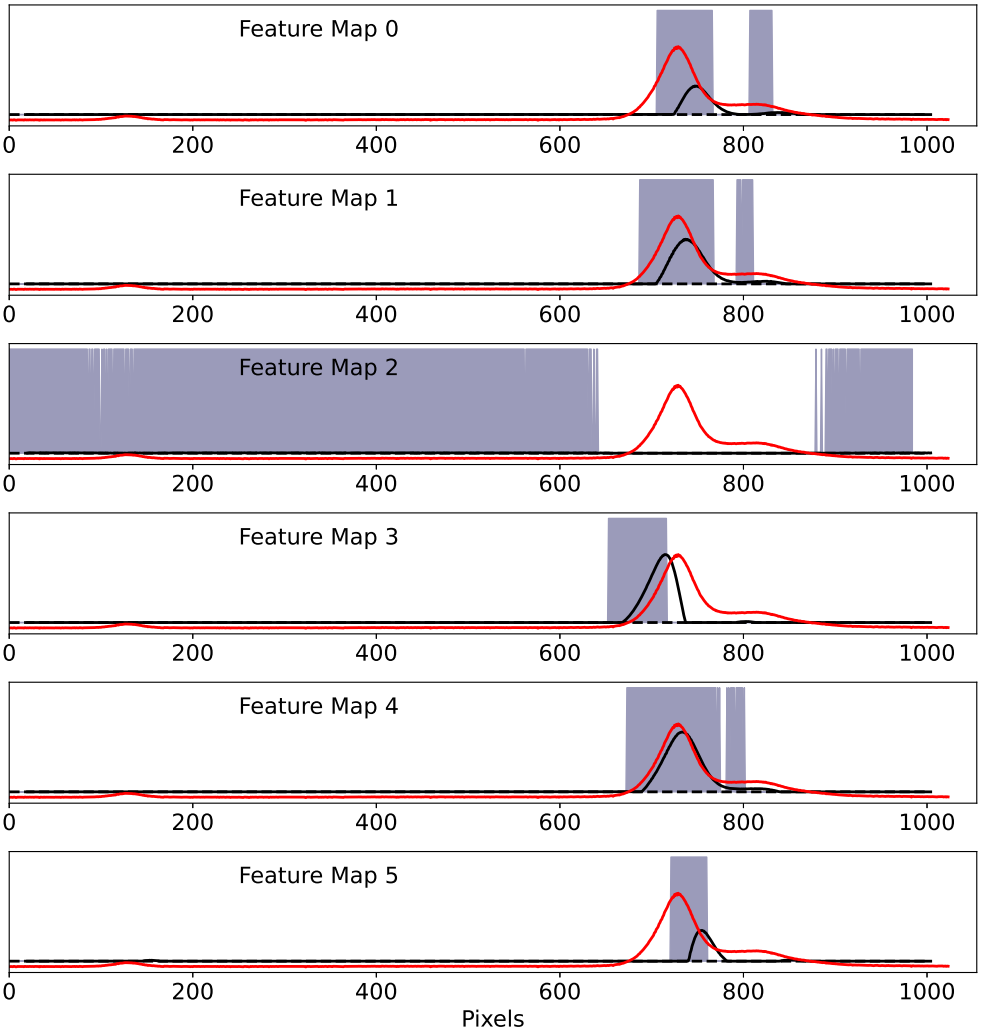

Figure 3. The 6 features maps generated by the first convolutional layer when applied to the D81 oil. All the feature maps (black line) has been normalised by dividing them by the maximum of the feature map 3 (the one with the largest maximum). After that all have been scaled so that Feature Map 3 has the same maximum as the measured spectra for readability. The input spectrum is indicated with red, the feature maps in black. The area shaded in purple indicate the pixels where the feature map value is greater than 5% of its maximum, or in other words. the areas where the feature maps are significantly different than zero.

D81 as it is easy to see that this discussion also applies to the feature maps obtained for the oil D03 (by only considering the correct feature map). The two peaks between pixel 700 and 850 are clearly the most important features. Feature Map 0, 1, 3, 4, and 5 are clearly related to these peaks. By looking at the black lines it is clear which part of the peaks is used as a feature. For example, Feature Map 3 and 5 identify the first peak increasing and decreasing parts respectively (the flanks of the peaks), although Feature Map 5 is smaller in values than Feature Map 3. Feature Map 2 identifies all other parts except the peaks, although its absolute value is much

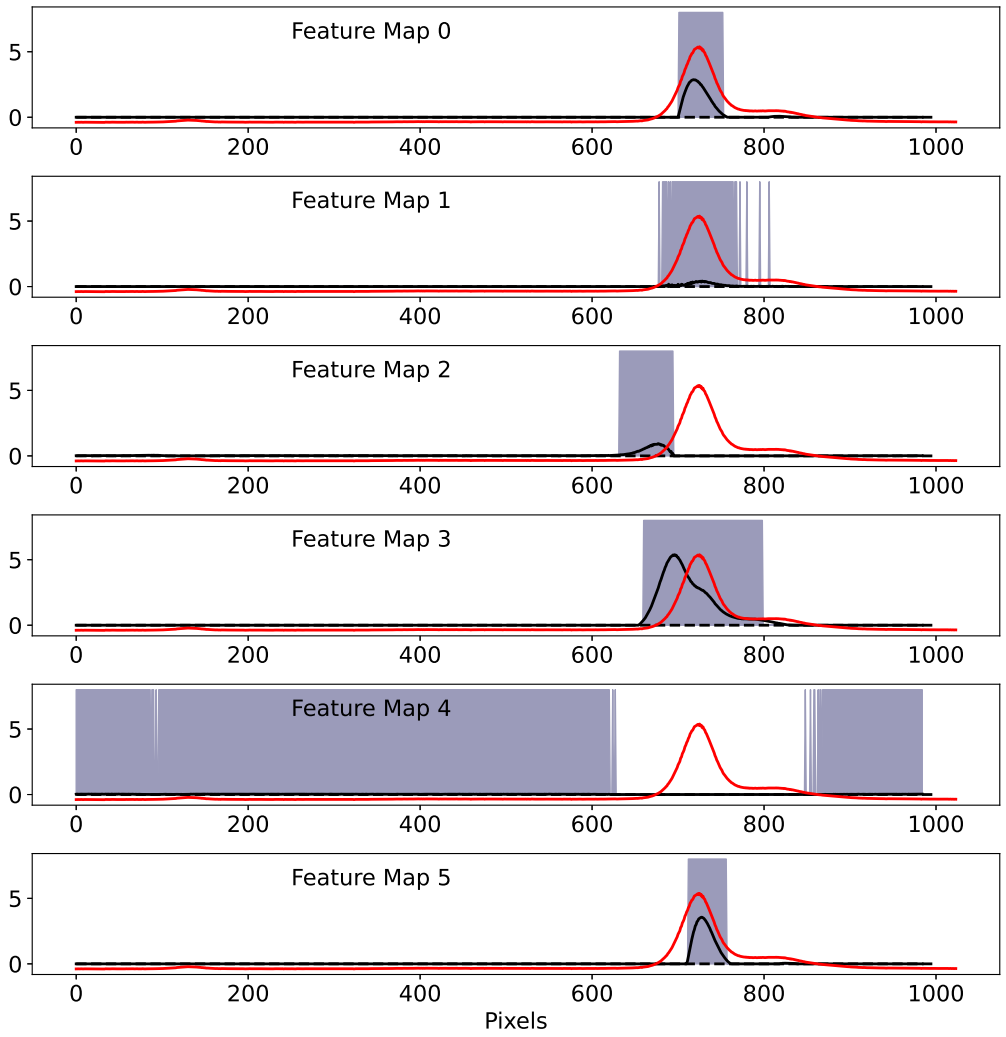

Figure 4. The 6 features maps generated by the first convolutional layer when applied to the D03 oil. All the feature maps (black line) has been normalised by dividing them by the maximum of the feature map 3 (the one with the largest maximum). After that all have been scaled so that Feature Map 3 has the same maximum as the measured spectra for readability. The input spectrum is indicated with red, the feature maps in black. The area shaded in purple indicate the pixels where the feature map value is greater than 5% of its maximum, or in other words, the areas where the feature maps are significantly different than zero.

smaller than the others relegating its importance to a clear second place with respect to the peaks. To give the reader a more quantitative way of estimating the ratios, the maximum of the Feature Map 0 is 44.8 times the one for Feature Map 2. In other words, Feature Map 2 is almost irrelevant for the prediction. Is nonetheless interesting to see, that the region outside the peaks is still identified as a somewhat interesting spectral window. Another very important result is that the peak at 397 nm (due to the excitation light of the LED and the optical geometry) is completely ignored by the network, as it contains no information on the chemical properties of the

olive oil. In this work, the feature maps of the second convolutional layer are not reported as their interpretation is less clear in fact they are the results of convolutional operations on the output of the max pooling layer that is composed of six different arrays therefore making their interpretation less useful.

For completeness, the metrics (the MSE) for the 1D-CNN used in this work are reported in Table 1. The values show that the network has learned quite effectively how to extract the parameter K_{232} for the two oils. Note that this is a clear pre-requisites for the feature maps to be studied. A network that has not learned to solve a given task, will of course also not identify reasonable features.

Oil	Dataset	MSE	MAE	MAE (%)
D81	Training	$6.9{\cdot}10^{-4}$	0.020	1%
D81	Validation	$1.3\cdot10^{-4}$	0.011	0.6%
D03	Training	$2.8{\cdot}10^{-4}$	0.013	0.7%
D03	Validation	$6.0\cdot10^{-4}$	0.020	1%

Table 1. Metric evaluated with the trained model on the training and validation dataset (composed of the oil D81). The MAE(%) is obtained by dividing the MAE value by the maximum value of the K_{232} parameter. The values show that the network has learned quite effectively how to extract the parameter K_{232}.

4. CONCLUSION

In conclusion, this work shows how, by studying feature maps obtained by a 1D-CNN, it is possible to identify the spectral regions of a measured spectrum that are related to specific chemical properties. This was demonstrated by applying it to the determination of the UV absorbance at 232 nm, K_{232}, from fluorescence spectra. The 1D-CNN successfully predicts the parameter K_{232} and enables, by studying feature maps, the clear identification of the relevant spectral features. Clearly, this method is not only interesting to explain how the 1D-CNN is using the input data, but it can be used to actually better understand the underlying chemical properties of substances. In this specific work olive oil has been used, but this can be applied to any other substances and, more generally, to any other type of spectra, making this approach a very interesting tool towards the understanding of optically active physico-chemical properties with the help of convolutional neural networks.

ACKNOWLEDGMENTS

This work was supported by the projects: "VIRTUOUS" funded by the European Union's Horizon 2020 Project H2020-MSCA-RISE-2019 Grant No. 872181; "SUSTAINABLE" funded by the European Union's Horizon 2020 Project H2020-MSCA-RISE-2020 Grant No. 101007702.

REFERENCES

[1] Yuanyuan, C. and Zhibin, W., "Quantitative analysis modeling of infrared spectroscopy based on ensemble convolutional neural networks," *Chemometrics and Intelligent Laboratory Systems* **181**, 1–10 (2018).

[2] Acquarelli, J., van Laarhoven, T., Gerretzen, J., Tran, T. N., Buydens, L. M., and Marchiori, E., "Convolutional neural networks for vibrational spectroscopic data analysis," *Analytica chimica acta* **954**, 22–31 (2017).

[3] Zhang, X., Lin, T., Xu, J., Luo, X., and Ying, Y., "Deepspectra: An end-to-end deep learning approach for quantitative spectral analysis," *Analytica chimica acta* **1058**, 48–57 (2019).

[4] Malek, S., Melgani, F., and Bazi, Y., "One-dimensional convolutional neural networks for spectroscopic signal regression," *Journal of Chemometrics* **32**(5), e2977 (2018).

[5] Ni, C., Wang, D., and Tao, Y., "Variable weighted convolutional neural network for the nitrogen content quantization of masson pine seedling leaves with near-infrared spectroscopy," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **209**, 32–39 (2019).

[6] Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B., "Definitions, methods, and applications in interpretable machine learning," *Proceedings of the National Academy of Sciences* **116**(44), 22071–22080 (2019).

[7] James, G., Witten, D., Hastie, T., and Tibshirani, R., [*An introduction to statistical learning*], vol. 112, Springer (2013).

[8] Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., and Engelsen, S. B., "Interval partial least-squares regression (i pls): A comparative chemometric study with an example from near-infrared spectroscopy," *Applied spectroscopy* **54**(3), 413–419 (2000).

[9] Arakawa, M., Yamashita, Y., and Funatsu, K., "Genetic algorithm-based wavelength selection method for spectral calibration," *Journal of Chemometrics* **25**(1), 10–19 (2011).

[10] Yun, Y.-H., Bin, J., Liu, D.-L., Xu, L., Yan, T.-L., Cao, D.-S., and Xu, Q.-S., "A hybrid variable selection strategy based on continuous shrinkage of variable space in multivariate calibration," *Analytica chimica acta* **1058**, 58–69 (2019).

[11] Zhang, J., Yan, H., Xiong, Y., Li, Q., and Min, S., "An ensemble variable selection method for vibrational spectroscopic data analysis," *RSC advances* **9**(12), 6708–6716 (2019).

[12] Bjerrum, E. J., Glahder, M., and Skov, T., "Data augmentation of spectral data for convolutional neural network (cnn) based deep chemometrics," *arXiv preprint arXiv:1710.01927* (2017).

[13] "Commission regulation (eec) no. 2568/91 of 11 july 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis official journal l 248, 5 september 1991," *Offic. JL* **248**, 1–83 (1991).

[14] Venturini, F., Sperti, M., Michelucci, U., Gucciardi, A., Martos, V. M., and Deriu, M. A., "Physico-chemical properties extraction from the fluorescence spectrum with 1d-convolutional neural networks: application to olive oil," *preprint (arXiv)* (2022).

[15] Venturini, F., Sperti, M., Michelucci, U., Herzig, I., Baumgartner, M., Caballero, J. P., Jimenez, A., and Deriu, M. A., "Exploration of spanish olive oil quality with a miniaturized low-cost fluorescence sensor and machine learning techniques," *Foods* **10**(5), 1010 (2021).

[16] Michelucci, U., [*Applied Deep Learning - A Case-Based Approach to Understanding Deep Neural Networks*], APRESS Media, LLC (2018).

[17] LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L., "Handwritten digit recognition with a back-propagation network," *Advances in neural information processing systems* **2** (1989).

[18] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE* **86**(11), 2278–2324 (1998).

[19] Michelucci, U., [*Advanced applied deep learning: convolutional neural networks and object detection*], APRESS Media, LLC (2019).

[20] Kingma, D.P.; Ba, J. A., "Adam: A method for stochastic optimization. in proceedings of 3rd.," *In Proceedings of 3rd International Conference on Learning Representations, ICLR 2015* , 1–15 (2015).