# Compact optical fluorescence sensor for food quality control using artificial neural networks: application to olive oil

# Compact optical fluorescence sensor for food quality control using artificial neural networks: application to olive oil

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

[a]TOELT llc, Machine Learning Research and Development, Birchlenstr. 25, 8600 Dübendorf, Switzerland
[b]Artificial Intelligence Laboratory, University of Ljubljana, Ljubljana, Slovenia
[c]Institute of Applied Mathematics and Physics, Zurich University of Applied Sciences, Technikumstrasse 9, 8401 Winterthur, Switzerland
[d]PolitoBIOMed Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy
[e]Department of Plant Physiology, Faculty of Sciences, Biotechnology Institute, Campus Fuentenueva s/n, 18071 University of Granada, Spain

## ABSTRACT

Olive oil is an important commodity in the world, and its demand has grown substantially in recent years. As of today, the determination of olive oil quality is based on both chemical analysis and organoleptic evaluation from specialized laboratories and panels of experts, thus resulting in a complex and time-consuming process. This work presents a new compact and low-cost sensor based on fluorescence spectroscopy and artificial neural networks that can perform olive oil quality assessment. The presented sensor has the advantage of being a portable, easy-to-use, and low-cost device, which works with undiluted samples, and without any pre-processing of data, thus simplifying the analysis to the maximum degree possible. Different artificial neural networks were analyzed and their performance compared. To deal with the heterogeneity in the samples, as producer or harvest year, a novel neural network architecture is presented, called here conditional convolutional neural network (Cond-CNN). The presented technology is demonstrated by analyzing olive oils of different quality levels and from different producers: extra virgin olive oil (EVOO), virgin olive oil (VOO), and lampante olive oil (LOO). The sensor classifies the oils in the three mentioned classes with an accuracy of 82%. These results indicate that the Cond-CNN applied to the data obtained with the low-cost luminescence sensor, can deal with a set of oils coming from multiple producers, and, therefore, showing quite heterogeneous chemical characteristics.

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

## 1. INTRODUCTION

Olive oil contains many healthy components,[1] like vitamins, and has beneficial effects like antioxidant activity.[2] Virgin olive oil, that is oil produced by the use of mechanical means only, is classified for commercial purposes in three decreasing quality grades: extra virgin olive oil (EVOO), virgin olive oil (VOO), and lampante olive oil (LOO). The best quality (EVOO) is of course the most expensive and beneficial, while the lowest grade (LOO) should not be consumed without refining according to the European Union regulations.[3,4] Therefore, the determination of the quality of olive oil is of the utmost importance. For this purpose, several approaches were proposed in the past including, for example, crystal microbalance arrays,[5] metal oxide sensors[6] and spectroscopic techniques in general.[7–9] Other approaches involve gas chromatography and mass spectrometry.[10,11]

Among the spectroscopic techniques, fluorescence spectroscopy is of great interest since it is a fast, cost-efficient, and at the same time sensitive method to study the properties of vegetables, particularly olive oils.[12,13]

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

The presence in olive oil of natural fluorescent molecules such as chlorophyll and beta-carotene, phenolic compounds, such as tocopherol, and their oxidation products make luminescence approaches very effective. The most common techniques are the acquisition of excitation-emission matrices (EEMs) or the use of synchronous scanning.[14] Applications include (non-exhaustive list): classification of different quality grades of olive oils,[15,16] adulteration detection,[17–19] monitoring of the oxidation processes,[7,20,21] deterioration monitoring,[22] and geographical origin determination.[23–25] The complexity in acquiring high-quality data (especially of EEMs) and in data post-processing limit largely the accessibility of these methods.

This work presents an easy and non-destructive method to determine the quality of olive oil from a simple single fluorescence spectrum. In particular, it describes an analysis of neural networks architectures to perform olive oil classification when dealing with 1-D spectra obtained with the low-cost sensor for fluorescence spectroscopy described by the authors in a previous publication.[26] The main contributions of this paper are three. Firstly, a large dataset composed of 45 different olive oils of three qualities (EVOO, VOO, and LOO) from two different producers and harvest years is presented. Secondly, three different neural network architectures are discussed and their performance in classifying olive oils' quality is compared. Thirdly, a new neural network architecture is proposed to deal with the heterogeneous chemical properties of oils, even when having the same declared quality.

## 2. MATERIALS AND METHODS

### 2.1 Oil Samples

In this study, 45 olive oils were investigated. 24 oils were provided by the producer Conde de Benalúa, Granada, southern Spain, from the 2019-2020 harvest and 21 from Producers of Eastern Mountains of the Province of Granada, southern Spain, from the 2020-2021 harvest. The oils are of three qualities: extra virgin olive oil (EVOO), virgin olive oil (VOO), and lampante olive oil (LOO). The quality assessment of all the olive oils was performed according to the current European Union regulations[3,4] for the commercialization of olive oil. An overview of the olive oils divided into the three quality classes and harvest year is reported in Table 1.

| Quality class | Number of samples | 2019-2020 harvest | 2020-2021 harvest |
|---|---|---|---|
| EVOO | 20 | 10 | 10 |
| VOO | 13 | 8 | 5 |
| LOO | 12 | 6 | 6 |

Table 1. Overview of the olive oils samples in each quality class. EVOO: extra virgin olive oil, VOO: virgin olive oil, LOO: lampante olive oil.

The oils of the two harvests have very different chemical and spectral characteristics, as it can be expected, being the oils produced from two different harvests and different producers, although all from the same geographical region. This heterogeneity of the oils poses a learning challenge for the neural networks because the spectra within one single quality class show very strong variations, sometimes stronger than between spectra from different classes.

### 2.2 Experimental

The fluorescence spectra were taken with the portable and low-cost sensor described in a previous publication.[26] The excitation light is provided by a UV LED with an emission at 395 nm. The fluorescence is collected by a miniature spectrometer from Ocean Optics placed at 90° with respect to the excitation LED. The resolution of the spectrometer is 16 nm. The spectrometer has a 1024-element CCD array which acquires the entire spectrum in one single measurement. The spectrum, therefore, consists of 1024 intensity values. The sensor can be seen in Figure 1 and was previously described in detail.[26]

All the measurements in this work were done on undiluted samples: for each oil, 4 ml were taken and placed in clear glass vials (as visible in Figure 1), and then measured 20 times. From the acquired spectra the background, or dark measurement, was subtracted, and the first 250 values, corresponding to wavelengths below 450 nm, were removed since they did not contain any fluorescence signal but may contain spurious stray light from the excitation LED. Finally, they were normalized to have an average of zero and a standard deviation of one.[27]

Figure 1. Photo of the fluorescence sensor an olive oil samples in the glass vials.

### 2.3 Machine Learning Models

In this work, three different types of architectures were investigated and compared to better understand the variability between the oils and to improve the prediction accuracy. The architectures are 1) a feed-forward neural network (FFNN), 2) a one-dimensional convolutional neural network (1D-CNN), and 3) a new type of architecture called here conditional convolutional neural network (Cond-CNN).

The input to the FFNN and 1D-CNN is the spectra after the background subtraction and normalization described in Section 2.2. Therefore, the input of the network is a one-dimensional row of 774 intensities. Both the FFNN and 1D-CNN do not take into account if the oil sample is coming from the single producer Conde de Benalúa, harvest 2019-2020, or from one of the other producers, harvest 2020-2021. Therefore, they have to try to extract features that are common to all samples, a task that may be very hard due to the heterogeneity discussed in Section 2.1.

The new Cond-CNN architecture proposed here addresses this issue by adding as input also a class: 0 if the sample is from Conde de Benalúa and 1 if coming from other producers. The additional input branch of the network is responsible for correcting the classification according to the producer of the input olive oil sample.

For all three neural network architectures, the cross-entropy loss function was chosen

$$L(\hat{y}_i, y_i) = -\frac{1}{N}\sum_{i=1}^{N}\left(y_i \log \hat{y}_i + (1-y_i)\log(1-y_i)\right) \tag{1}$$

where $y_i$ is the true expected class for the $i^{th}$ input spectra, $\hat{y}_i$ is the predicted class for the $i^{th}$ input spectra from the network, $N$ is the size of the dataset used.

To evaluate the performance of the neural networks two additional metrics were monitored: the accuracy $a$ and the confusion matrix $C$.[27] Defining the identity function $\mathbb{I}(x, y)$ as

$$\mathbb{I}(x, y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

the accuracy $a$ over a dataset of size $N$ can be defined as

$$a = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(\hat{y}_i, y_i). \tag{3}$$

The confusion matrix $C \in \mathbb{R}^{3\times3}$ can be also defined with the help of the identity function as

$$C_{k,l} = \sum_{i=1}^{N} \mathbb{I}(\hat{y}_i, l-1)\mathbb{I}(y_i, k-1) \tag{4}$$

where $l \in \{1,2,3\}$ and $k \in \{1,2,3\}$. The metrics reported in the results section are the mean of the accuracy over the 10 splits indicated with $\langle a \rangle$ and its standard deviation over the splits, indicated with $\sigma(a)$.

To better judge the generalization capacity of the trained neural networks the dataset was split to include the spectra of 35 oils (*training dataset*) and those of 10 oils (*validation dataset*). Each model was trained on the training dataset and its performance evaluated on the validation dataset, therefore, the validation is done on spectra of unseen oils. To avoid overfitting an early-stopping strategy was adopted.[27] The model that gave the best performance on the validation dataset during the training was then saved and used. This process was repeated 10 times (effectively performing 10 random splits) to be able to estimate the variance of the results on different dataset splits. The number 10 was chosen as a compromise between running time and confidence of the estimation. In the results section, the averaged confusion matrix over the 10 splits is reported, with the values given in percent of the total number of inputs in the validation dataset.

For all the three neural networks, to deal with the light unbalance of the data, the loss function was minimized by using class weights inversely proportional to the respective class frequencies.[28,29] In all three cases the network optimizer chosen was Adam.[30] Note that overfitting was carefully monitored. In particular, by using early stopping and drop-out,[27] the architectures were optimised until the validation accuracy and training accuracy were very close, meaning that the network outputs the same accuracy results on the training set and on the validation test. In particular, all the results described here were obtained with models where the training accuracy and validation accuracy are well within one standard deviation of each other. The details relevant to each of the different architectures are given in Sections 2.3.1 to 2.3.3.

### 2.3.1 Feed Forward Neural Network Architecture

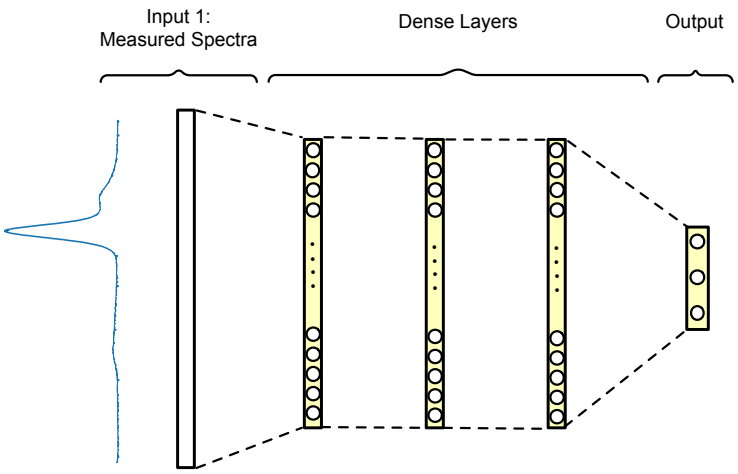


Figure 2. Architecture of the FFNN used in this work. The yellow elements symbolize dense layers.

The architecture of the FFNN network used in this work is depicted in Figure 2, and consists of three layers, each having 32 neurons. The ReLU activation function was chosen. A drop-out rate of 0.1 in each dense layer was used. A mini-batch size of 64 was used and the network was trained for 500 epochs.

### 2.3.2 1D-CNN Architecture

The 1D-CNN network, depicted in Figure 3, is composed by the following sequence of layers:

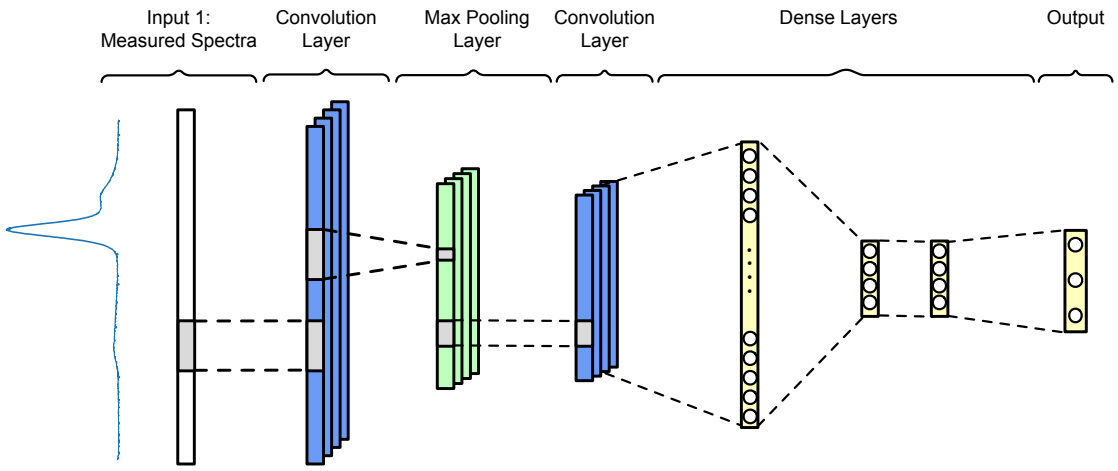


Figure 3. Architecture of the 1D-CNN used in this work. The yellow elements symbolize dense layers, the blue elements symbolize convolutional layers, the green elements symbolyze max pooling layer.

1. convolutional layer with 6 filters each of size 40;
2. max pooling layer with a size of 10;
3. convolutional layer with 6 filters each of size 20 (with a drop-out layer with a drop-out rate of 0.5);
4. a dense layer obtained flattening and concatenating the output of the previous layer;
5. dense layer with 4 neurons;
6. dense layer with 4 neurons;
7. dense layer with 3 neurons (output layer).

Also for this architecture the ReLU activation function[27] was chosen. A mini-batch size of 64 was used and the network was trained for 500 epochs.

### 2.3.3 Conditional 1D-CNN Architecture

The new network architecture proposed to address the heterogeneity of the oils is depicted in Figure 4. It is characterized by two inputs: the measured spectra and the information of the producer or harvest year of the oil. Note that this second input, here called producer class, is categorical, thus in the Python implementation its value has been one-hot encoded.[27] In other words, effectively the second input is a tuple that is $(1, 0)$ for Conte de Benalúa oils of the 2019-2020 harvest and $(0, 1)$ for all others oils.

The network used in this paper is composed by the following sequence of layers for respectively the top and bottom branches in Figure 4:

1. top branch in Fig. 4 - convolutional layer with 6 filters each of size 20;
2. top branch in Fig. 4 - max pooling layer with a size of 10;
3. top branch in Fig. 4 - convolutional layer with 6 filters each of size 10;
4. top branch in Fig. 4 - a dense layer obtained flattening and concatenating the output of the previous layer;
5. top branch in Fig. 4 - dense layer with 32 neurons;
6. top branch in Fig. 4 - dense layer with 16 neurons;

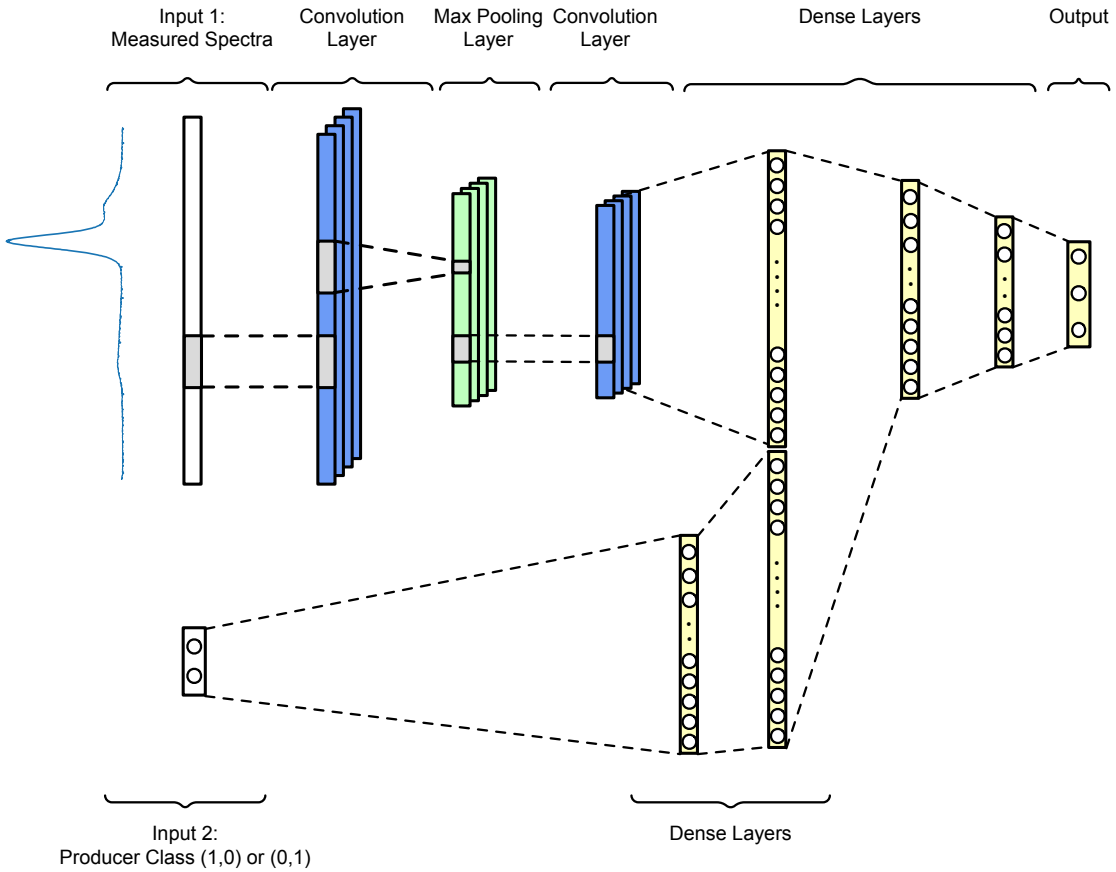


Figure 4. Architecture of the Cond-CNN used in this work. The yellow elements symbolize dense layers, the blue elements symbolize convolutional layers, the green elements symbolyze max pooling layer.

7. bottom branch in Fig. 4 - dense layer with 16 neurons;
8. bottom branch in Fig. 4 - dense layer with 36 neurons;
9. dense layer with 3 neurons (output layer).

For all the layers the swish activation function[31] was used. For the results presented in this paper a mini-batch size of 8 was used. The models were all trained for 3000 epochs. In this case no drop-out was used as early stopping was enough to avoid overfitting.

## 3. RESULTS AND DISCUSSION

The three ANN described in Section 2.3 were tested by predicting the quality of unseen olive oils performing a classification into the three classes EVOO, VOO, and LOO. The results illustrated with the confusion matrix are shown in Figure 5. The confusion matrices present the accuracy of each network averaged over ten different splits of the dataset and reported as % of the number of oil in each class.

Figure 5 shows that the performance of the Cond-CNN exceeds that of the two other networks. In particular, the Con-CNN can distinguish exceptionally well between edible (EVOO and VOO) and non-edible (LOO) oils. Only 1% of the LOO oils were misclassified as VOO and only 2% of EVOO and 2% of VOO were wrongly classified as LOO. The result is reasonable since the quality of the EVOO and VOO may be very close and differ, for example, in the organoleptic properties, which are difficult to extract from a single fluorescence spectra.

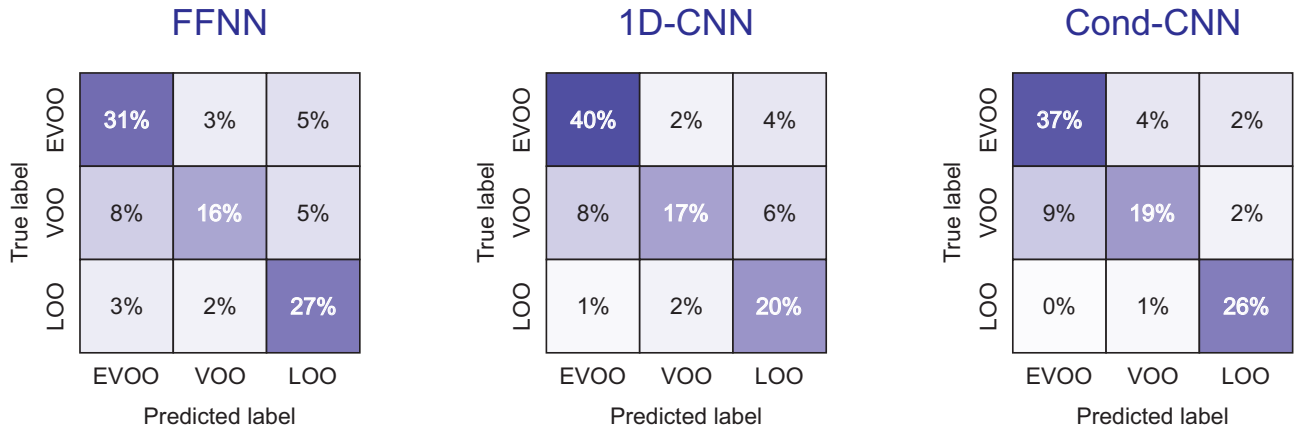


Figure 5. Confusion matrix visualising the accuracy in the prediction of the quality of olive oil using the three types of neural networks discussed in this work. The value reported are given as % of the oils in each class and are averaged over 10 splits.

The performance is further illustrated by the calculation of the accuracy $a$ defined in Equation 3. The results for each of the ANN architectures are reported in Table 2. $\langle a \rangle$ indicates the mean of the accuracy, and $\sigma(a)$ its standard deviation when evaluated over the validation dataset for 10 splits.

| **Network Architecture** | $\langle a \rangle$ | $\sigma(a)$ |
|---|---|---|
| FFNN | 0.74 | 0.10 |
| 1D-CNN | 0.77 | 0.09 |
| Cond-CNN | 0.82 | 0.06 |

Table 2. Performance comparison the different neural network architectures. $a$ indicates the accuracy, $\langle a \rangle$ its mean, and $\sigma(a)$ its standard deviation evaluated over 10 different splits.

The inferior performance of the FFNN and of the 1D-CNN is not surprising and can be understood in terms of the heterogeneity of the oils. These two architectures do not take into account the different years of harvest and different producers and, therefore, assume that all the 45 olive oils have similar and comparable characteristics. The results indicate that this is not the case. Adding the information that there are two groups of oils (one would probably need more than just two producer classes), the use of the Cond-CNN architecture increases the. Due to the lack of specific information (e.g. cultivar, producer) of the 21 oils of the 2020-2021 harvest, it is not possible to obtain a much higher accuracy than 82% obtained with the Cond-CNN architecture.

The high standard deviation is also related to the small number of oils used for the validation (10 oils). Since the accuracy is an averaging metric, the central limit theorem (CLT)[32] states that the smallest the sample size, the largest the variance (and thus the standard deviation) of an averaging metric. It is possible to estimate what is the expected standard deviation of the accuracy by using statistical arguments. In fact, one can note that the accuracy, as written in Equation 3, is the average of a Bernoulli variable ($\mathbb{I}(\hat{y}_i, y_i)$) with a probability $p \approx a$ of being one. Thus, the average over $n$ samples of the above-mentioned Bernoulli variables will follow a Binomial distribution. For binomial distribution the central limit theorem starts to be a reasonable approximation for values of $np = na > 10$[33] . In our case $na \approx 7$ for FFNN and $na \approx 8$ for the Cond-CNN. In this case, the condition $np = na > 10$ is not really satisfied and an estimation from the CLT can only be considered a rough approximation, but one that can let us understand if the results that were obtained are reasonable. It is easy to show that, defining $I_i = \mathbb{I}(\hat{y}_i, y_i)$, one can write $\mathrm{Var}(I_i) = \sigma^2(I_i) = a - a^2$ therefore making the calculation of the $\mathrm{Var}(I_i)$ very easy. Due to the small number of samples the variance should be calculated by using the Bessel's correction[33] by using $n-1$ in the denominator for the calculation of the variance instead of $n$. The CLT tells* that the average of a sample of size $n$ of the $I_i$ (the accuracy) will have a standard deviation that when $n \to \infty$ will tend to $\sigma/\sqrt{n}$.[33] With the sample size that has been used here $n = 10$, the CLT gives an estimate

*The CLT validity condition will not be discussed here, although it can be easily verified that they are satisfied.

of the standard deviation of the accuracy of the order of 0.1 (the exact values for the three cases reported in Table 2 are 0.13, 0.12 and 0.11 respectively). This matches the order of magnitude of the obtained results and the fact that the standard deviation diminishes with increasing accuracy. The differences between this estimate and the observed results are due to three main factors: the small sample size (therefore the CLT is not a good approximation), the fact that $a$ and, therefore, $p$ are only approximation evaluated with a sample size of 10, and the fact that $\sigma$ is also a rough approximation evaluated with a sample size of 10 (although corrected with the Bessel's correction). In conclusion, such a large standard deviation is not a reflection of the model being unstable, but more of the small sample size.

## 4. CONCLUSIONS

This work presents a new compact and low-cost sensor based on fluorescence spectroscopy and artificial neural networks that can perform olive oil quality assessment. Thanks to an innovative low-cost sensor and a novel neural network architecture, the conditional one-dimensional convolutional neural network (Cond-CNN), it is possible to perform a classification of olive oil into the three quality classes: EVOO, VOO, and LOO. The method uses a portable, easy-to-use, and low-cost device, which works with undiluted samples, and without any pre-processing and is, therefore, fast, simple to use, and cost-effective.

The performance of the new Cond-CNN architecture exceeds that of the feed-forward neural network and of the one-dimensional convolutional neural network since can take into account the heterogeneity of the spectral and chemical characteristics of the oils in the dataset. The accuracy obtained with this sensor is 82%. Most of the errors in the classification occur for the distinction of EVOO and VOO, which probably have very close chemical characteristics and may differ for properties not significantly affecting the fluorescence spectrum. If the sensor is used to distinguish edible (EVOO and VOO) from non-edible (LOO) oils, the accuracy reaches 95%.

These results show still a large potential of fluorescence spectroscopy when combined with artificial neural networks for the assessment of the quality of food and food technology in general.

## ACKNOWLEDGMENTS

This work was supported by the projects: Innosuisse - Swiss Innovation Agency, Grant No. 36761.1 INNO-LS; "SUSTAINABLE" funded by the European Union's Horizon 2020 H2020-MSCA-RISE-2020 program, Grant No. 101007702; "PARENT" funded by the European Union's Horizon 2020 H2020-MSCA-ITN-2020 program, Grant No. 956394; "Project of Excellence" from Junta de Andalucia-FEDER-Fondo de Desarrollo Europeo 2018. Ref. P18-H0-4700.

## REFERENCES

[1] Willett, W. C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E., and Trichopoulos, D., "Mediterranean diet pyramid: a cultural model for healthy eating," *The American journal of clinical nutrition* **61**(6), 1402S–1406S (1995).

[2] Kris-Etherton, P. M., Hecker, K. D., Bonanome, A., Coval, S. M., Binkoski, A. E., Hilpert, K. F., Griel, A. E., and Etherton, T. D., "Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer," *The American journal of medicine* **113**(9), 71–88 (2002).

[3] "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).

[4] "Commission implementing regulation no 1348/2013 of december 17 2013," *Official Journal of the European Union* **338**, 31–67 (2013).

[5] Escuderos, M. E., Sánchez, S., and Jiménez, A., "Quartz crystal microbalance (qcm) sensor arrays selection for olive oil sensory evaluation," *Food Chemistry* **124**(3), 857–862 (2011).

[6] García-González, D. L. and Aparicio, R., "Classification of different quality virgin olive oils by metal-oxide sensors," *European Food Research and Technology* **218**(5), 484–487 (2004).

[7] Baltazar, P., Hernández-Sánchez, N., Diezma, B., and Lleó, L., "Development of rapid extra virgin olive oil quality assessment procedures based on spectroscopic techniques," *Agronomy* **10**(1), 41 (2020).

[8] Violino, S., Ortenzi, L., Antonucci, F., Pallottino, F., Benincasa, C., Figorilli, S., and Costa, C., "An artificial intelligence approach for italian evoo origin traceability through an open source iot spectrometer," *Foods* **9**(6), 834 (2020).

[9] Martín-Tornero, E., Fernández, A., Pérez-Rodriguez, J. M., Durán-Merás, I., Prieto, M. H., and Martín-Vertedor, D., "Non-destructive fluorescence spectroscopy as a tool for discriminating between olive oils according to agronomic practices and for assessing quality parameters," *Food Analytical Methods* , 1–13 (2021).

[10] Ruiz-Samblás, C., Cuadros-Rodríguez, L., González-Casado, A., de Paula Rodríguez García, F., de la Mata-Espinosa, P., and Bosque-Sendra, J. M., "Multivariate analysis of ht/gc-(it) ms chromatographic profiles of triacylglycerol for classification of olive oil varieties," *Analytical and bioanalytical chemistry* **399**(6), 2093–2103 (2011).

[11] Taiti, C., Marone, E., Fiorino, P., and Mancuso, S., "The olive oil dilemma: To be or not to be evoo? chemometric analysis to grade virgin olive oils using 792 fingerprints from ptr-tof-ms," *Food Control* , 108817 (2022).

[12] Kongbonga, Y. G. M., Ghalila, H., Onana, M. B., Majdi, Y., Lakhdar, Z. B., Mezlini, H., and Sevestre-Ghalila, S., "Characterization of vegetable oils by fluorescence spectroscopy," *Food and Nutrition Sciences* **2**(7), 692–699 (2011).

[13] Sikorska, E., Khmelinskii, I., and Sikorski, M., "Analysis of olive oils by fluorescence spectroscopy: methods and applications," *Olive oil-constituents, quality, health properties and bioconversions* , 63–88 (2012).

[14] Skoog, D. A., Holler, F. J., and Crouch, S. R., [*Principles of instrumental analysis*], Cengage learning (2017).

[15] Guimet, F., Boqué, R., and Ferré, J., "Cluster analysis applied to the exploratory analysis of commercial spanish olive oils by means of excitation- emission fluorescence spectroscopy," *Journal of agricultural and food chemistry* **52**(22), 6673–6679 (2004).

[16] Poulli, K. I., Mousdis, G. A., and Georgiou, C. A., "Classification of edible and lampante virgin olive oil based on synchronous fluorescence and total luminescence spectroscopy," *Analytica Chimica Acta* **542**(2), 151–156 (2005).

[17] Sayago, A., Morales, M., and Aparicio, R., "Detection of hazelnut oil in virgin olive oil by a spectrofluorimetric method," *European Food Research and Technology* **218**(5), 480–483 (2004).

[18] Poulli, K. I., Mousdis, G. A., and Georgiou, C. A., "Rapid synchronous fluorescence method for virgin olive oil adulteration assessment," *Food chemistry* **105**(1), 369–375 (2007).

[19] Ali, H., Saleem, M., Anser, M. R., Khan, S., Ullah, R., and Bilal, M., "Validation of fluorescence spectroscopy to detect adulteration of edible oil in extra virgin olive oil (evoo) by applying chemometrics," *Applied spectroscopy* **72**(9), 1371–1379 (2018).

[20] Hernández-Sánchez, N., Lleó, L., Ammari, F., Cuadrado, T. R., and Roger, J. M., "Fast fluorescence spectroscopy methodology to monitor the evolution of extra virgin olive oils under illumination," *Food and Bioprocess Technology* **10**(5), 949–961 (2017).

[21] Mishra, P., Lleó, L., Cuadrado, T., Ruiz-Altisent, M., and Hernández-Sánchez, N., "Monitoring oxidation changes in commercial extra virgin olive oils with fluorescence spectroscopy-based prototype," *European Food Research and Technology* **244**(3), 565–575 (2018).

[22] Lobo-Prieto, A., Tena, N., Aparicio-Ruiz, R., García-González, D. L., and Sikorska, E., "Monitoring virgin olive oil shelf-life by fluorescence spectroscopy and sensory characteristics: A multidimensional study carried out under simulated market conditions," *Foods* **9**(12), 1846 (2020).

[23] Dupuy, N., Le Dréau, Y., Ollivier, D., Artaud, J., Pinatel, C., and Kister, J., "Origin of french virgin olive oil registered designation of origins predicted by chemometric analysis of synchronous excitation- emission fluorescence spectra," *Journal of agricultural and food chemistry* **53**(24), 9361–9368 (2005).

[24] Jiménez-Carvelo, A. M., Lozano, V. A., and Olivieri, A. C., "Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity confirmation and geographical origin of argentinean extra virgin olive oils," *Food Control* **96**, 22–28 (2019).

[25] Al Riza, D. F., Kondo, N., Rotich, V. K., Perone, C., and Giametta, F., "Cultivar and geographical origin authentication of italian extra virgin olive oil using front-face fluorescence spectroscopy and chemometrics," *Food Control* **121**, 107604 (2021).

[26] 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).

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

[28] Wu, Z., Guo, Y., Lin, W., Yu, S., and Ji, Y., "A weighted deep representation learning model for imbalanced fault diagnosis in cyber-physical systems," *Sensors* **18**(4), 1096 (2018).

[29] Aurelio, Y. S., de Almeida, G. M., de Castro, C. L., and Braga, A. P., "Learning from imbalanced data sets with weighted cross-entropy function," *Neural processing letters* **50**(2), 1937–1949 (2019).

[30] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," *Proceedings of 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA* , 1–15 (2015).

[31] Ramachandran, P., Zoph, B., and Le, Q. V., "Searching for activation functions," *arXiv preprint arXiv:1710.05941* (2017).

[32] Michelucci, U. and Venturini, F., "Estimating neural network's performance with bootstrap: A tutorial," *Machine Learning and Knowledge Extraction* **3**(2), 357–373 (2021).

[33] Casella, G. and Berger, R. L., [*Statistical inference*], Cengage Learning (2021).