Measuring fruit quality traits in olive through RGB imaging and artificial neural networks: opportunities and limitations

Giuseppe Montanaro
DiCEM
University of Basilicata
Potenza, Italy
giuseppe.montanaro@unibas.it

Angelo Petrozza
Centro Ricerche Metapontum Agrobios
ALSIA Metaponto di Bernalda (MT), Italy
angelo.petrozza@alsia.it

Laura Rustioni
DiSTeBA
University of Salento
Lecce, Italy
laura.rustioni@unisalento.it

Francesco Cellini
Centro Ricerche Metapontum Agrobios
ALSIA Metaponto di Bernalda (MT), Italy
francesco.cellini@alsia.it

Antonio Carlomagno
DiCEM
University of Basilicata
Potenza, Italy
antonio.carlomagno@unibas.it

Vitale Nuzzo
DiCEM
University of Basilicata
Potenza, Italy
vitale.nuzzo@unibas.it

***Abstract*—Imaging is an emerging contact-less high-throughput technology employed to retrieve quantitative and qualitative plant traits. In addition, it is often combined with artificial neural networks (ANNs) to further improve the reliability of image-based digital proxies. The olive oil industry is expanding globally as olive oil is increasingly recognized as a functional food. Fast and reliable determination of fruit quality traits is challenging in the agricultural sector. This study summarizes recent advances in the use of RGB-based imaging combined with ANNs to (i) predict oil and phenol concentrations in olive fruit and (ii) classify fruit at harvest according to colour and defects. Opportunities and limitations are also discussed.**

Keywords—olive oil, classification, RGB images, *Olea europaea*, polyphenols, regression.

I. Introduction

Within the digital agricultural domain, image-based methods are employed to retrieve several plant traits enabling the estimation of water status, incidence of diseases, fruit defects at harvest, and growth (e.g., [1], [2], [3], [4]). The use of imaging to identify proxies of quality traits during the fruit growing season is developing, too [5]. In addition, in the agricultural sector, as in others, imaging is often combined with artificial neural networks (ANNs) models to solve prediction and classification problems.

The olive oil industry is expanding globally as olive oil is increasingly recognized as a functional food [5]. Within the olive sector, ANNs have been employed mainly for characterizing olive oil and analyzing its composition. For example, [4] reviewed the use of predictive methods based on ANN with application during the production and processing of olive oil. Although the quality of olive oil substantially depends on the quality of olives, the application of imaging and ANNs during the field stage of olive fruit development received little attention.

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022, CUP: C33C22000250001). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

The quality of raw olives (e.g., maturity level, absence of external defects) is pivotal for the production of high quality table olives and olive oil. For this purpose, the use of ANN and imaging are increasingly used to support fast and reliable classification and sorting of raw olives [3][7].

This study summarizes recent achievements in imaging olive fruit as combined with ANNs focusing on both the prediction of fruit quality during development and the classification of raw olives at harvest. Opportunities and limitations are also discussed.

II. Description of the Contribution

In the olive sector, as in most horticultural sectors, fruit quality traits should be clearly defined to better support their accomplishment. In the case of olives, fruit quality is essential for olive oil determination. Particularly, phenolic compounds are substantial quality traits of olive oil not only for their healthy nutritional attributes but also for other technological properties, including contrasting oil oxidation throughout the storage time [8]. Hence, harvest time should be decided to maximize the content of phenols. However, the amount of extractable oil is still the dominant trait determining the profitability of olive crop. Considering that both oil and phenol concentrations in olive fruit accumulate during field cultivation, the decision of harvest time becomes relevant for olive fruit and, in turn, for olive oil quality. Following this, both phenols and oil concentrations need to be monitored to identify the best harvest time.

For table olives, the absence of damages and uniformity of colour is associated with high consumers' acceptance of the fruit phenotype [8][9]. The absence of defects in fruit is also a prerequisite for high-quality olive oil, including its stability during storage [10]. Hence, a process aimed at classifying and sorting olives is recommended for top productions.

Based on own papers and those from the literature, the present contribution highlights current advances and critical issues on RGB-based imaging and ANNs methods aimed at (i) predicting the change of key fruit quality traits (oil and phenols concentrations) during fruit development, and (ii) to classifying olive fruit at harvest according to colour and eventual defects.

III. Results and Applicability

Prediction of fruit quality development. Examples of the concentration of total polyphenols and oil in developing olive fruit throughout the season are reported in Fig. 1. It could be noted that oil concentration accumulated almost linearly, while total polyphenols had a fast early increase thereafter, they remain roughly stable. Accumulation patterns and absolute values of polyphenols and oil are influenced by the cultivar suggesting the training of any model should be tailored to a specific cultivar and would cover the whole developmental period.

During development, fruit phenotype showed the classical change of skin colour (Fig. 2). However, the colour veraison process would depend on the cultivar or at least on the ripening group (i.e., early, medium, or late ripening cultivar) [11]. Hence, Within an ANNs context with an ideal ANN working simultaneously over several cultivars, the ripening period of the cultivar(s) under scrutiny should be provided as a piece of known information to an to improve model performance [11].

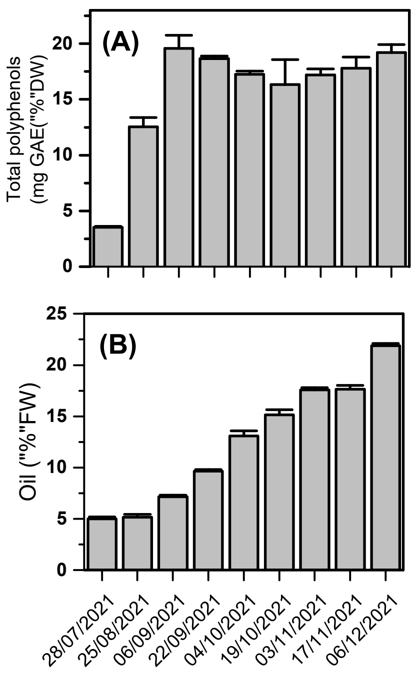

Fig. 1. Concentration of (A) total polyphenols (mg Gallic Acid Equivalent per mg of Dry Weight –flesh and stone), and (B) oil. Bars are ±SE.

The phenotype variation could be well described as changes in R, G, and B mean pixel values. For example, in the Leccino cultivar, these values showed an initial increase in every color band, peaking at maximum values in early September (Fig. 3). Thereafter, the values of R, G, and B color bands declined toward the minimum values achieved at the beginning of November. However, R and G promptly declined to values of approx. 66% and 80% of the maximum, respectively. By contrast, the mean pixel value of the B color band had a more gentle and progressive decline of about 20% of the maximum (Fig. 3). Using R, G and B clearly improves a classical visual colourimetric parameter used for olive ripening class assignment (i.e. Jaèn's index) which suffers poor standardization due to variable *visus* of personnel [3].

The seasonal dynamics of fruit phenotype (i.e., skin colour) in various cropping areas might differentially be coupled with that of oil and polyphenols because of a variable *gene* × *environment* interaction. Hence, appraisal of the influence of different environment on the phenotype of the same cultivar needs to be specifically tested within an imaging context.

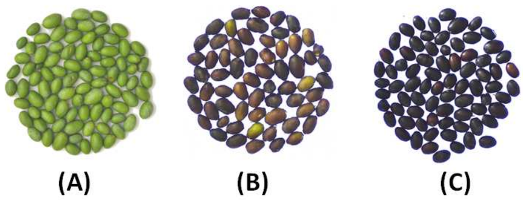

Fig. 2. Example of phenotype of "Leccino" olive samples collected on (A) end of July, (B) beginning of September and (C) mid October.

Within ANNs, the Back Propagation Neural Network model (BPNN) fed with RGB-based colourimetric indexes, had a good performance in predicting the oil concentration over a dataset covering the whole developmental stage (coefficient of determination R^2 = 0.94) (Fig. 4A). As concern the polyphenols, the BPNN showed a weaker prediction performance compared to that of oil (R^2 =0.79) (Fig. 4B). This lower performance can be explained considering that the imaging only collected the variations of R, G and B colour bands of the skin, while polyphenol variations would also be associated to the change of colour of the flesh layer according to the abovementioned common colorimetric Jaèn's index [3]. Therefore, in the case of polyphenols, the model integration with signals able to collect deep layer of fruit (e.g., spectral) would be substantial to improve its accuracy.

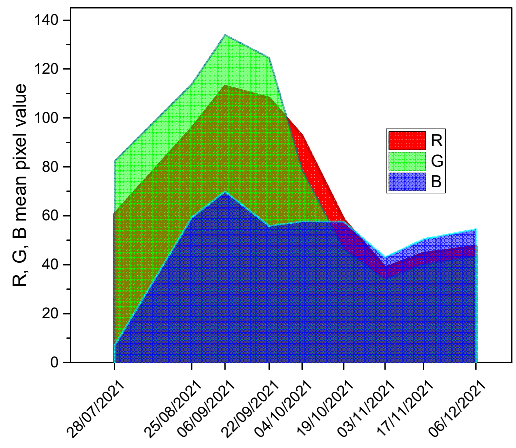

Fig. 3. Variation of the mean pixel values of the red (R), green (G) and blue (B) color channels measured in olive samples througout the growing season. Redrwan from [11].

The combination of image-based phenotyping and an ANN model to predict the concentration of oil and polyphenols in olive fruit creates the basis for further research aimed at large scale application of the method to assist growers' decision on optimal harvest time. In addition, the use of RGB sensors would be in line with affordable phenotyping as these sensors are cheap and easily accessible (e.g., through mobile phones) [12]. However, a large exploitation of results requires more effort to test the model in different cultivars having various ripening times and different absolute levels of oil and polyphenol concentrations. In addition, as a common issue in plant phenotyping, efforts are required to move on from lab to field applications accounting for variable environmental conditions which might influence the standardization of RGB signals.

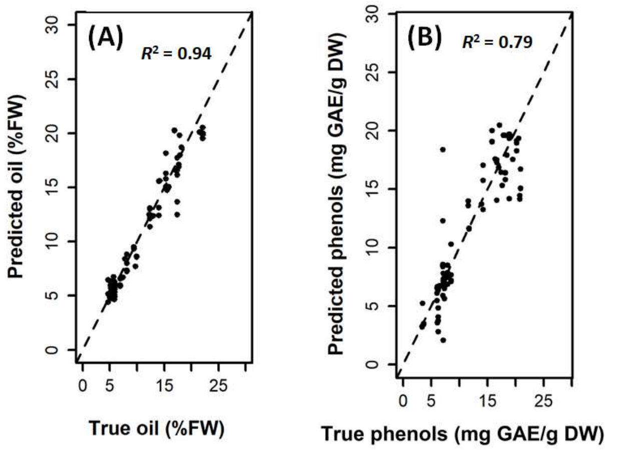

Fig. 4. Correlation between measured and predicted (A) oil and (B) polyphenols concentrations extracted from developing olive fruit. Predictions were achieved through a BPNN with a single hidden layer and 35 RGB-based colourimetric indexes as input variables. Figures have been produced based on source data redrawn from [11].

<u>Fruit classification and sorting at harvest.</u> Olive fruit for table consumption requires uniformity of the maturity level which might be associated with fruit phenotype. In addition, in table olives fruit external damages influence consumers' acceptance [8][10]. Defects of fruit also reduce the quality of the olive oil and, in turn, its shelf-life [10]. Hence, fruit classification and sorting before processing is recommended for top productions.

Following this, the combination of RGB-based imaging and ANN is employed for automatic and high-throughput classification of fruit for both oil and table olives production [10]. Particularly, Convolutional Neural Network modeling (YOLO) is used for the selection of fruit based on skin colour and eventual defects [5][7][8][10].

Recent studies have focused on various fruit defects, including the presence of discoloured spots, mechanical and biotic damages, and integrity [7][10]. In a laboratory study, using a high-resolution RGB camera, the resulting classification accuracy of an ANN model was above 95% for both colour classification and defects identification [10]. Similarly, a model based on RGB imaging coupled with a K-nearest neighbors algorithm [8] was able to explain about 90% of the total variance of the colour of olives (Fig. 5). Within an automated pipeline for detection of defects, the speed of the algorithm execution is a substantial feature of the model. However, a trade-off between algorithm execution time and efficacy in object detection is often targeted in classification problems [10].

In conclusion, recent scientific achievements provide substantial progress in using RGB-based imaging and ANN within regression and classification problems. Future efforts are required to solve criticisms related to cultivar-specific issues (ripening time) and to the influence of the environment. In addition, limitations of RGB (e.g., lack of information on deep flesh layer) should be overcome by the combination of different sensors (e.g., VIS–NIR, UV, hyperspectral, and fluorescence imaging).

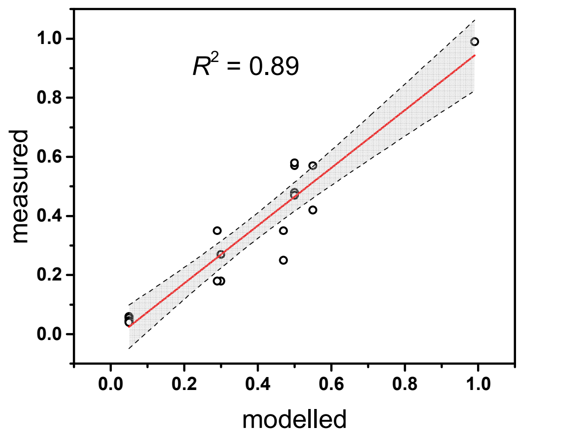

Fig. 5. Correlation between measured and modelled fraction (%) of black olives using RGB images of trayes and the k-nearest neighbors (k-NN) clustering algorithm for pixel classification. The grey area represents the confidence interval around the linear model (solid line). Elaboration by authors on source data appearing in [8].

REFERENCES

[1] Briglia N., Williams K., Wu D., Li Y., Tao S., Corke F., Montanaro G., Petrozza A., Amato D., Cellini F., Doonan J.H., Yang W. and V. Nuzzo, "Image-based assessment of drought response in grapevines", Front. Plant Sci. 11:595, 2020.

[2] Briglia N., Montanaro G., Petrozza A., Summerer S., Cellini F., and V. Nuzzo. "Drought phenotyping in Vitis vinifera using RGB and NIR imaging", Scientia Horticulturae, 256, DOI 10.1016/j.scienta.2019.108555, 2019.

[3] R. Furferi, L. Governi, and Y. Volpe, "ANN-based method for olive Ripening Index automatic prediction" Journal of Food Engineering, vol. 101, pp. 318–328, 2010.

[4] I., Gonzalez-Fernandez, M. A. Iglesias-Otero, M. Esteki, O. A. Moldes, J. C. Mejuto, and J. Simal-Gandara. "A Critical Review on the Use of Artificial Neural Networks in Olive Oil Production, Characterization and Authentication", Critical Reviews in Food Science and Nutrition 59(12), 1913–26, 2019.

[5] A. Aprile, C. Negro, E. Sabella, A. Luvisi, F. Nicolì, E. Nutricati, M. Vergine, A. Miceli, F. Blando, and L. De Bellis, "Antioxidant Activity and Anthocyanin Contents in Olives (Cv Cellina Di Nardò) during Ripening and after Fermentation," Antioxidants vol. 8, no. 5, pp. 138, 2019 https://doi.org/10.3390/antiox8050138

[6] M. D. Navarro-Hortal, J. M. Romero-Márquez, V. Jiménez-Trigo, J. Xiao, F. Giampieri, T. Y. Forbes-Hernández, G. Grosso, M. Battino,

C. Sánchez-González and J. L. Quiles, “Molecular bases for the use of functional foods in the management of healthy aging: Berries, curcumin, virgin olive oil and honey; three realities and a promise,” Critical Reviews in Food Science and Nutrition, 2022.

[7] S. Figorilli, S. Violino, L. Moscovini, L. Ortenzi, G. Salvucci, S. Vasta, F. Tocci, C. Costa, P. Toscano, and F. Pallottino, “Olive fruit selection through ai algorithms and RGB imaging,” Foods, vol. 11, p. 3391, 2022 https://doi.org/10.3390/foods11213391

[8] F. Gutiérrez, T. Arnaud and A. Garrido, “Contribution of polyphenols to the oxidative stability of virgin olive oil,” Journal of the Science of Food and Agriculture, vol. 81, no. 15, pp. 1463-1470, 2001 https://doi.org/10.1002/jsfa.958

[9] L. Ortenzi et al., ‘A Machine Vision Rapid Method to Determine the Ripeness Degree of Olive Lots’, Sensors, vol. 21, no. 9, Art. no. 9, Jan. 2021, doi: 10.3390/s21092940.

[10] G. Salvucci, ‘Fast olive quality assessment through RGB images and advanced convolutional neural network modeling’, European Food Research and Technology, p. 11, 2022.

[11] G. Montanaro, A. Petrozza, L. Rustioni, F. Cellini, V. Nuzzo, “ Phenotyping key fruit quality traits in olive using RGB images and back propagation neural networks”, Plant Phenomics, 5:Article 0061. https://doi.org/10.34133/plantphenomics.0061.

[12] D. Reynolds, F. Baret, C. Welcker, A. Bostrom, J. Ball, J., F. Cellini, F., A. Lorence, A., Chawade, M. Khafif, K. Noshita, M. Mueller-Linow, J. Zhou, F. and Tardieu, F. , ”What is cost-efficient phenotyping? Optimizing costs for different scenarios,” Plant Science, vol. 282, pp. 14-22, 2019.