Artificial Intelligence to Identify Olive-Tree Cultivars

João Mendes[1], José Lima[1], Lino Costa[2], and Ana I. Pereira[1]

[1] Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal {joao.cmendes, jllima, apereira}@ipb.pt
[2] ALGORITMI Center, University of Minho, 4710-057 Braga, Portugal
lac@dps.uminho.pt

Abstract. The exponential advance in artificial intelligence techniques makes it possible to apply them to previously thought to be impossible sectors. In this work, a different approach is presented to identify the different varieties of olive trees present in the olive groves of Portugal. Using its leaves and deep learning algorithms necessary for its classification, the proposed system can perform a reliable, low-cost, and real-time identification of the olive trees.

Keywords: Artificial intelligence · Olive · Classification.

1 Introduction

The production of olives, whether for olive oil or table olives, is a growing business, according to the Observatory of Economic Complexity (OEC), in 2020, approximately 1.23 billion dollars were handled in world trade, with Portugal being the third largest exporter with approximately 6.9% of the total amount [1].

In Portugal, olive growing is an activity of high economic, social and environmental importance. There are 361 483 hectares of olive groves in the country, in about 118 450 farms [12]. The olive tree is distributed from north to south of the country, with particular incidence in the interior regions, with Alentejo, Trás-os-Montes, Beira Interior and Ribatejo e Oeste being the main producing regions. In the last campaign (2021), provisional data from the National Statistics Institute point to a record harvest with an expected oil production of 2 25 million hectoliters [13]. This production is obtained through the various olive-tree cultivars present in our country, whether autochthonous or not. These blends of varieties are often beneficial as they each have unique chemical, physical and flavor characteristics [10,18] that are commonly used either to improve the season or date of harvest through the selection of varieties with different maturation periods, or to ensure a more harmonious composition of the oil from an organoleptic point of view.

In this way, and highlighting the notorious chemical and physical differences of the different olive-tree cultivars, it is extremely important for the olive oil producer to carry out a reliable and rapid identification of the varieties present in their olive groves. In this sense, a bibliographic research was carried out on the various identification techniques applied over time. Two terms, "Olive identification cultivars" or "Olive variety identification" were searched in two of the best-known databases (Web of Science (WOS) and Scopus). The research resulted in 914 and 534 documents in the WOS and Scopus databases, respectively, after their collection, the R software was used, combined with the bibliometrix tool [2], where trends and research techniques were analyzed as well as the main keyword groupings.

Through the results obtained, it was possible to perceive that most of the techniques applied for the identification of varieties consist of genetic analysis, techniques such as random amplified polymorphic DNA (RAPD), molecular markers, microsatellites, chemometrics and DNA Fingerprints [3–5, 7, 8, 20]. Techniques that have high reliability of responses, but are somewhat time-consuming processes that do not allow us to identify them on the spot, with the need to collect samples to analyze in the laboratory and with a relevant associated cost when one considers that in the same olive grove there can be dozens of different species. Other techniques related to identification through artificial vision have also been studied and implemented with encouraging hit rates [6, 17, 21]. On the other hand, these techniques always use the tree's fruit to proceed with its classification, thus being restricted to a specific time of the year to proceed with its identification.

Analyzing the problems that arise from the techniques presented above, the focus of this work is to ensure a form of identification on the spot, with minimal impact on the tree, and that can be carried out at any season of the year. In this way and analyzing what is being done in other cultures, the solution presented involves the use of artificial intelligence algorithms to identify the leaves of the tree. This is a solution that solves practically all the problems shown above, guarantees an instant classification without having to resort to specialized technicians, it is possible to do it every year, since the olive tree is a permanent leaf tree. It also guarantee that the impact for the tree and the cost associated with the process is null.

This implementation becomes easier taking into account the advances in the area of artificial intelligence algorithms, following the hardware developments. This type of algorithms has proven to be highly qualified for solving this type of tasks, being applied to various species such as pistachios [11], grapes [16, 19], apples [15] and other examples [9, 14].

In this way, the proposed approach is summarized by the use of artificial intelligence algorithms to identify the different autochthonous olive tree species.

This papers is organized as follows. After the introduction, section 2 presents the work methodology where the system architecture is approached and the focus of this article is highlighted. Finally, the third chapter concludes the article and points out future work.

2 Methodology

As the main objective of this work is the development of an artificial intelligence system capable of identifying the different varieties of olive-tree cultivars. For that, there is a set of steps to be taken and implemented, thus ensuring the correct functioning of the system. Figure 1 presents the system architecture.

As a common need for all artificial intelligence problems, the dataset is undoubtedly the most important element of these implementations. In this way, and since it is an innovative approach to the problem in question, it was necessary to proceed with the creation of a sufficiently large dataset of images to guarantee a good behavior of the classification algorithms. To this end, and with the help of partners specialized in the area, several visits were made to olive groves to collect the necessary material. It was

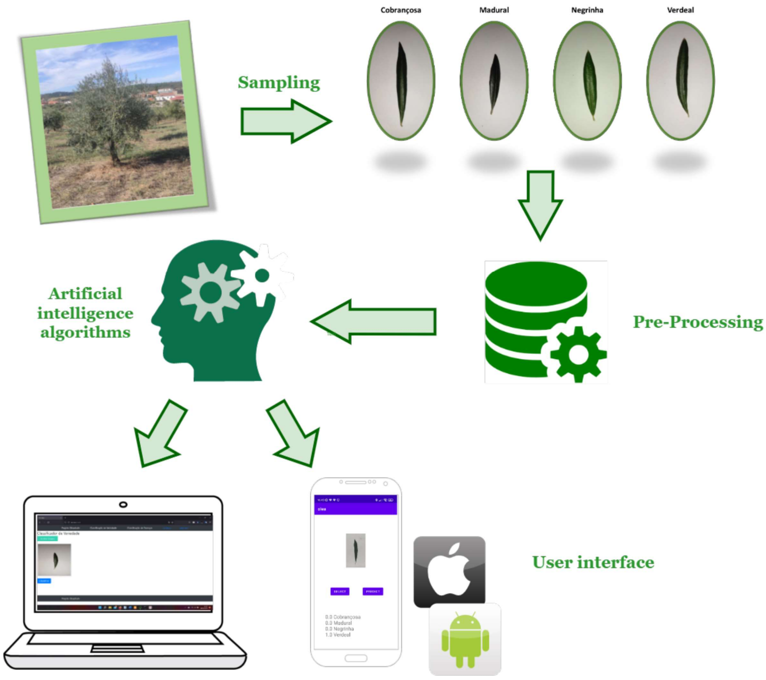

Fig. 1: System architecture.

stipulated at the outset to make collections in the different seasons of the year in order to accompany the growth and recession of the leaves, as well as other factors that have an impact on their development, such as water or nutrient stress, thus ensuring a correct identification regardless of the phase and conditions in which the tree is found.

After each sample collection, several processes need to be carried out, starting with the photographic capture of each sample on a white background to facilitate the algorithm training process, then there is a pre-processing of the images in order to reduce their size, optimizing the functioning of the algorithms. In this pre-processing, the photos were cut through autonomous processes, resized to a resolution of 299x299 pixels, and further normalized in its three layers (RGB). Once these processes are completed, it will be possible to implement the artificial intelligence algorithms. This implementation will consist of three steps: training the algorithm using the training set, classifying the test data, and analyzing it. After the implementation and the adjustments, the models will be made available on two platforms, the web version, and the mobile version, thus allowing their use by all interested stakeholders, whether they are consumers, producers, or just enthusiasts in the area.

3 Conclusion and future work

Constant hardware improvements are enabling a significant advance in the use of artificial intelligence methods for the most diverse purposes. As it was possible to prove with the cited references, there are already some works that propose similar approaches in different cultures, demonstrating the possibility of success of the presented study.

As the main future works of the study, it is possible to highlight the continuation of updating the dataset with more samples and the application of various machine and deep learning methods to enable the best possible results. It culminates in an easy-to-use application that aims to make it possible to identify on the spot, without additional costs for the user, the different varieties present in the national territory.

References

1. The observatory of economic complexity. `https://oec.world/en/profile/hs/olive-oil-fractions-refined-not-chemically-modifie`, accessed: 2022-05-10
2. Aria, M., Cuccurullo, C.: bibliometrix: An r-tool for comprehensive science mapping analysis. Journal of informetrics **11**(4), 959–975 (2017)
3. Bautista, R., Crespillo, R., Cánovas, F.M., Gonzalo Claros, M.: Identification of olive-tree cultivars with scar markers. Euphytica **129**(1), 33–41 (2003)
4. Besnard, G., Baradat, P., Bervillé, A.: Olive cultivar identification using nuclear rapds and mitochondrial rflps. In: International Symposium on Molecular Markers for Characterizing Genotypes and Identifying Cultivars in Horticulture 546. pp. 317–324 (2000)
5. Besnard, G., Breton, C., Baradat, P., Khadari, B., Bervillé, A.: Cultivar identification in olive based on rapd markers. Journal of the American Society for Horticultural Science **126**(6), 668–675 (2001)
6. Beyaz, A., Özkaya, M.T., İçen, D.: Identification of some spanish olive cultivars using image processing techniques. Scientia Horticulturae **225**, 286–292 (2017)
7. Ergulen, E., Ozkaya, M., Ulger, S., Ozilbey, N.: Identification of some turkish olive cultivars by using rapd-pcr technique. In: IV International Symposium on Olive Growing 586. pp. 91–95 (2000)
8. Fabbri, A., Hormaza, J., Polito, V.: Random amplified polymorphic dna analysis of olive (olea europaea l.) cultivars. Journal of the American Society for Horticultural Science **120**(3), 538–542 (1995)
9. Grinblat, G.L., Uzal, L.C., Larese, M.G., Granitto, P.M.: Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture **127**, 418–424 (2016)
10. Guinda, A., Lanzón, A., Albi, T.: Differences in hydrocarbons of virgin olive oils obtained from several olive varieties. Journal of Agricultural and Food Chemistry **44**(7), 1723–1726 (1996)
11. Heidary-Sharifabad, A., Zarchi, M.S., Emadi, S., Zarei, G.: An efficient deep learning model for cultivar identification of a pistachio tree. British Food Journal (2021)
12. INE: Instituto nacional de estatística, estatísticas agrícolas de base. `https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_base_dados`, accessed: 2022-05-11
13. INE: Instituto nacional de estatística, previsões agrícolas. `www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=526211517&DESTAQUESmodo=2`, accessed: 2022-05-11
14. Larese, M.G., Granitto, P.M.: Hybrid consensus learning for legume species and cultivars classification. In: European Conference on Computer Vision. pp. 201–214. Springer (2014)
15. Liu, C., Han, J., Chen, B., Mao, J., Xue, Z., Li, S.: A novel identification method for apple (malus domestica borkh.) cultivars based on a deep convolutional neural network with leaf image input. Symmetry **12**(2), 217 (2020)
16. Liu, Y., Su, J., Shen, L., Lu, N., Fang, Y., Liu, F., Song, Y., Su, B.: Development of a mobile application for identification of grapevine (vitis vinifera l.) cultivars via deep learning. International Journal of Agricultural and Biological Engineering **14**(5), 172–179 (2021)
17. Martínez, S.S., Gila, D.M., Beyaz, A., Ortega, J.G., García, J.G.: A computer vision approach based on endocarp features for the identification of olive cultivars. Computers and Electronics in Agriculture **154**, 341–346 (2018)
18. Montaño, A., Sánchez, A., Casado, F., De Castro, A., Rejano, L.: Chemical profile of industrially fermented green olives of different varieties. Food Chemistry **82**(2), 297–302 (2003)
19. Nasiri, A., Taheri-Garavand, A., Fanourakis, D., Zhang, Y.D., Nikoloudakis, N.: Automated grapevine cultivar identification via leaf imaging and deep convolutional neural networks: a proof-of-concept study employing primary iranian varieties. Plants **10**(8), 1628 (2021)
20. Shahriari, M., Omrani, A., Falahati-Anbaran, M., Ghareyazei, B., Nankali, A.: Identification of iranian olive cultivars by using rapd and microsatellite markers. In: V International Symposium on Olive Growing 791. pp. 109–115 (2004)

21. Vanloot, P., Bertrand, D., Pinatel, C., Artaud, J., Dupuy, N.: Artificial vision and chemometrics analyses of olive stones for varietal identification of five french cultivars. Computers and Electronics in Agriculture **102**, 98–105 (2014)