ABSTRACT

The purpose of this study is to analyze the capability of Ag/Si-wafer Surface-Enhanced Raman Spectroscopy (SERS) substrates in distinguishing pure EVOO spectra from other vegetable oils. This study also aims to quantify variance sources across the hierarchical structure of SERS measurements, as well as to evaluate the performance of machine learning in differentiating pure EVOO from adulterated EVOO. The dataset comprised 2,460 spectra from 145 vials, covering EVOO Bertolli and five adulterant types at concentrations of 1-50% v/v. Spectra were acquired using a portable Raman spectrometer Brolight with Maya 2000 Pro detector from Ocean Optics at 785 nm excitation. The results show that the six pure oil classes exhibit high spectral similarity, with dominant variance originating from Ag/Si-wafer substrate heterogeneity. Hierarchical variance decomposition quantified inter-sample differences as the largest variance contributor, which is systematic and reproducible. The machine learning model 100% successfully distinguished pure EVOO from adulterated EVOO on an independent test set limited to 24 vials. Model decisions were confirmed to be grounded in chemical signals rather than substrate artifacts. This claim remains limited to the Bertolli EVOO brand with a relatively small dataset size, so generalization to other EVOO types requires further validation.

Keywords**:** *surface-enhanced Raman spectroscopy, extra virgin olive oil, hierarchical variance decomposition, support vector machine, feature importance analysis*