Research Article

Journal of Environmental Nanotechnology

Quantitative Detection and Over-limit Identification of Sesamol, Octyl Gallate and Ferulic Acid in Soybean Oil Using AgNP-based SERS and Supervised Learning Algorithms

M. Gomathy[1], A. Vijaya Lakshmi[2], M. Prabu[3], V. Vignesh[4], N. Nirmal Singh[5] and A. Jose Anand[6*]

[1]Department of Computer Applications, SGT University, Gurugram, HR, India
[2]Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Shamshabad, Hyderabad, TG, India
[3]Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Kuthambakkam, Tiruvallur, TN, India
[4]Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, TN, India
[5]Department: of Electronics and Communication Engineering, DMI College of Engineering, Chennai, TN, India
[6]Department of Electronics and Communication Engineering, KCG College of Technology, Karapakkam, Chennai, TN, India
Received: 23.02.2026 Accepted: 11.03.2026 Published: 30.03.2026
*joseanand97@gmail.com

ABSTRACT

The excessive or uncontrolled addition of synthetic and natural antioxidants in edible oils raises increasing food safety concerns and necessitates rapid and sensitive detection techniques. In this study, a surface-enhanced Raman spectroscopy (SERS) platform integrated with machine learning algorithms was developed for the detection and discrimination of sesamol (SM), octyl gallate (OG), and ferulic acid (FA) in soybean oil. Silver nanoparticles (AgNPs) with an average size of 15–20 nm were synthesized using a citrate reduction method and employed as SERS substrates. The synthesized AgNPs exhibited a strong surface plasmon resonance peak at 528 nm, a high enhancement factor of 4.4×10^5, and good reproducibility, with a relative standard deviation (RSD) of 8.17%. Quantitative analysis showed a strong linear relationship between SERS peak intensity and antioxidant concentration in the range of 50–400 mg/kg, with correlation coefficients of 0.9383 (SM), 0.9761 (OG), and 0.9607 (FA). The limits of detection were 9.711 mg/kg for SM, 8.107 mg/kg for OG, and 7.345 mg/kg for FA. To further enhance identification accuracy, several supervised machine learning models, including support vector machine (SVM), random forest (RF), and k-nearest neighbors (KNN), were applied to classify SERS spectral data. Among these models, SVM achieved the best classification performance with an accuracy of 97.11%, while over-limit detection at 200 mg/kg showed an overall accuracy greater than 93%, including 100% prediction accuracy for FA and SM samples. The integration of SERS with machine learning provides a rapid, sensitive, and reliable analytical strategy for food safety monitoring and quality control of antioxidants in edible oils, demonstrating strong potential for real-time regulatory screening applications.

Keywords: Antioxidants; Random Forest; Octyl gallate; Nanoparticles; Citrate reduction.

1. INTRODUCTION

Edible oils are necessary for human diets, supplying energy, essential fatty acids, and fat-soluble vitamins, all critical for appropriate growth and development. Octyl gallate (OG) is a widely used synthetic phenolic antioxidant that improves oxidative stability in edible oils. Previous studies have reported its effectiveness in inhibiting lipid oxidation and enhancing oil shelf life under various processing conditions. However, excessive addition of such antioxidants may pose safety concerns, highlighting the need for rapid and reliable detection methods (Avuloglu-Yilmaz *et al.* 2023). Mechanistically, OG was identified as a critical chain-length antioxidant in water-in-oil (W/O) emulsions, exhibiting dual interfacial action that enhances radical collecting efficiency at the oil–water interface (Zhang *et al.* 2023). Furthermore, the classical "cut-off effect" and "polarity paradox" were experimentally validated for OG and related gallates in bulk oils and emulsified systems, thereby reinforcing their structure–activity relationship in lipid oxidation control (Pei *et al.* 2022). Comprehensive reviews have also highlighted environmental and health considerations associated with synthetic antioxidants, emphasizing the need for advanced monitoring strategies (Wang *et al.* 2021). Practical studies in frying oil and seafood systems confirmed that OG and related gallates significantly improved oxidative stability while adhering to polarity and cut-off principles (Li *et al.* 2021). In polyunsaturated systems such as docosahexaenoic acid DHA oil, synergistic combinations of OG with other antioxidants markedly prolonged shelf life (Shen *et al.* 2021).

Natural antioxidants such as ferulic acid (FA) have attracted attention as sustainable alternatives to

synthetic additives due to their strong radical scavenging activity and potential applications in edible oil stabilization (Rodríguez-Blázquez *et al.* 2024). Sunflower hull-derived FA and associated phenolics exhibited strong radical scavenging activity suitable for oil preservation and active packaging applications (Nobili *et al.* 2021). Gurum seed oil extracted through screw press, supercritical fluid extraction (SFE), and solvent extraction methods showed substantial phenolic content, including FA, confirming its richness in natural antioxidants (Karrar *et al.* 2020). Similarly, halophytic seed oils containing FA, gallic acid, and flavonoids demonstrated strong antioxidant performance, highlighting their potential as sustainable antioxidant sources in edible oils (Toqeer *et al.* 2018).

Among naturally occurring antioxidants, sesamol (SM) is another important natural phenolic antioxidant known for its strong oxidative stability and effectiveness under thermal processing conditions in the edible oil system (Ahmad *et al.* 2023). A switchable deep eutectic solvent system incorporating sesamol enabled selective extraction and monitoring of tert-butylhydroquinone (TBHQ) in edible oils with high HPLC-UV sensitivity (Liu *et al.* 2022). In deep-frying systems, SM significantly suppressed total polar component formation and improved metabolic health indicators (Wang *et al.* 2022). Thermal transformation studies revealed that sesamol forms a dimer in soybean oil with superior antioxidant performance, confirmed through HPLC-MS and NMR analysis (Chen *et al.* 2022). Additionally, sesamol effectively inhibited glycidyl ester formation in palm oil during deep frying and deodorization, showing comparable or superior performance to TBHQ at optimized concentrations (Han *et al.* 2021; Han *et al.* 2022). Under thermal stress conditions, SM significantly extended the induction periods and reduced oxidation rate constants in olive, sesame, and canola oils (Toorani *et al.* 2021). Microwave-assisted synthesis further enhanced sesamyl butyrate formation, improving antioxidant stability in oil systems (Keramat and Golmakani, 2020).

To ensure rapid, sensitive, and real-time monitoring of antioxidants in edible oils, SERS platforms integrated with machine learning have emerged as powerful analytical tools. Silver nanoparticle (AgNP)-based SERS systems, validated through UV–Vis spectroscopy and microscopy, have demonstrated strong potential for trace detection of phenolic antioxidants (Kamil *et al.* 2017). Advanced nanostructured substrates such as COFs-Au@AgNPs enabled ultra-trace detection of contaminants in olive oil using deep learning algorithms (Wu *et al.* 2025). Similarly, TTN@AgNP substrates achieved ppb-level detection in edible oils (Dhillon *et al.* 2025), while Ag/Cu nanozyme-based SERS selectively detected gallate antioxidants with improved sensitivity (Zhang *et al.* 2025). Ag@Au hybrid SERS substrates coupled with artificial neural networks (ANN) successfully predicted peroxide values and fatty acid profiles with high accuracy (Nagpal *et al.* 2023). Silver-based SERS approaches have also demonstrated nanomolar-level detection capability for phthalates in edible oils (Wang *et al.* 2023b).

Conventional techniques such as chromatographic methods are widely used for detecting antioxidants in edible oils; however, they are often time-consuming, require complex sample preparation, and are not suitable for rapid on-site monitoring. Therefore, there is a need for fast and sensitive analytical methods for antioxidant detection in food systems. In this study, a surface-enhanced Raman spectroscopy (SERS) platform integrated with machine learning algorithms was developed for identifying antioxidants in edible oils. AgNPs synthesized via a citrate reduction method were used as SERS substrates for detection within a 50–400 mg/kg concentration range. Several machine learning models, including RF, KNN, SVM, stacked autoencoder, and CNN, were evaluated for antioxidant classification, while PLS-DA (partial least squares–discriminant analysis) was used for qualitative discrimination. The novelty of this work lies in the integration of SERS with machine learning to enable rapid and reliable monitoring of antioxidants in edible oils for food safety applications.

2. MATERIALS AND METHODS

2.1 Materials

Sesamol (SM, $C_7H_6O_3$), Octyl gallate (OG, $C_{15}H_{22}O_5$), and Ferulic acid (FA, $C_{10}H_{10}O_4$) were utilized as antioxidant ligands. Silver nitrate ($AgNO_3$) served as the silver precursor for synthesizing silver nanoparticles (AgNPs), whereas trisodium citrate dihydrate ($Na_3C_6H_5O_7 \cdot 2H_2O$) acted as both the reducing and stabilizing agent. Methylene blue (MB, $C_{16}H_{18}ClN_3S$) served as the probe molecule. Methanol (CH_3OH) and anhydrous ethanol (C_2H_5OH) were used as analytical-grade solvents. Soybean oil was utilized as the primary medium for the assessment of antioxidants. Distilled and deionized water were utilized in all studies to guarantee solution purity, avert ionic interference, and sustain experimental repeatability.

2.2 Synthesis of AgNPs

Silver nanoparticles (AgNPs) measuring around 15–20 nm were manufactured by a chemical reduction technique utilizing sodium citrate as reducing and stabilizing agent and silver nitrate as precursor. First, 17 mg of $AgNO_3$ was dissolved in 100 ml of deionized water and heated to boiling under continuous magnetic stirring. After reaching boiling temperature, 2 mL of 1% (w/v) sodium citrate solution was rapidly added to the reaction mixture. The solution was maintained at boiling for 30–40 min, during which the color changed from colorless to yellowish-brown, indicating nanoparticle formation

(Assad *et al.* 2026). Upon completion of the reaction, the colloidal solution was permitted to cool naturally to ambient temperature while stirring persisted. The produced AgNPs were subsequently placed in centrifuge tubes and spun at 8000 rpm for 10 min to eliminate surplus reactants. The supernatant was removed, and the precipitated nanoparticles were reconstituted in deionized water. The obtained AgNPs were redispersed in deionized water and stored at 4 °C in dark conditions until further use. All synthesis experiments were repeated five times to ensure reproducibility.

2.3 UV–Vis and Raman Spectroscopy

The optical properties and performance of the substrates were assessed using a T9s UV–Vis spectrophotometer (Shimadzu UV-2600), with absorbance measurements taken from a cuvette filled with deionized water as the blank reference. Raman spectra were obtained utilizing an Andor Shamrock i-Raman series spectrometer.

2.4 Experimental Sample Classification

The samples designated for the trials were classified into three groups based on the distinct experimental aims. The initial batch was formulated to assess the effectiveness of the SERS substrate, utilizing MB solutions at concentrations spanning from 10^{-3} to 10^{-9} M, made with distilled water (Mohanasundaram *et al.* 2024). The second batch consisted of high-concentration ethanol solutions of SM, OG, FA, and their binary combinations (OG + FA, SM + OG, and SM + FA) to investigate the spectrum properties of different antioxidants. The third batch comprised six strengths of contaminated oil samples formulated to evaluate the antioxidant content in genuine oil systems containing a singular antioxidant.

2.5 SERS Measurement and Pretreatment Protocol

The SERS spectra of the reference solutions were obtained using a confocal Raman microscope (Andor Shamrock i-Raman series). Before measurement, the pertinent acquisition parameters were configured as follows: excitation wavelength of 532 nm, 10× objective for initial focusing, 50× objective for spectral collection, laser power of 25 mW, 600 g/mm grating, integration period of 10 s, and three accumulations per spectrum. Spectral calibration was executed utilizing the distinctive Raman peak of single-crystal silicon at 511 cm^{-1} to rectify any deviations induced by slight optical misalignments. In SERS analysis, 5 µL of standard solution was thoroughly combined with an equivalent volume of AgNPs solution, and 5 µL of the resultant mixture was drop-cast onto an ultrasonically cleaned single-crystal silicon substrate. To mitigate the coffee-ring effect, substrates were permitted to rest for 10 s post-deposition, followed by a two-step spin-coating procedure at 1000 rpm for 10 s and 3000 rpm for 30 s to create a homogeneous thin film. Initial alignment was conducted with the 10× long-working objective to calibrate focus and stage height, then utilizing a 50× long-focus objective for spectrum capture. The Raman spectra were shown and acquired in real time utilizing the instrument software.

A pretreatment technique was implemented on the antioxidant-infused oil samples before analysis due to the intricate composition of edible oils. Precisely, 6 g of the oil sample with a specified antioxidant content was combined with 3.5 mL of methanol and subjected to magnetic stirring for 4 min to guarantee thorough extraction of the antioxidant into the methanol phase. To achieve phase separation, the mixture was centrifuged at 7,000 rpm for 12 min. The methanolic layer was meticulously recovered, while the remaining oil underwent 2 more extraction procedures to optimize antioxidant recovery. Aggregated methanol extracts were subsequently employed for SERS detection of SM. Each oil sample was prepared and analyzed in triplicate, and the obtained SERS spectra were averaged to ensure experimental reliability and reproducibility. The pretreatment oil extracts were subsequently evaluated using the confocal Raman microscope system in accordance with the previously outlined SERS measurement protocol.

The obtained spectrum data included details regarding the chemical composition and concentration of the materials, as well as extraneous factors like noise and cosmic ray interference. Moreover, discrepancies in sample composition and characteristics may influence the observations, resulting in baseline drift and spectrum changes. For each sample, five independent measurements were performed at different positions on the substrate, and the obtained spectra were averaged to improve reliability. This procedure ensured experimental reproducibility and minimized signal variation caused by substrate heterogeneity.

2.6 Description of Proposed Algorithm Framework

2.6.1 Principal Component Analysis (PCA)

Large datasets are frequently produced to aid decision-making; yet, their increasing complexity presents substantial obstacles for effective analysis and administration. Dimensionality reduction is crucial for enhancing data management and interpretability. Principal Component Analysis (PCA) is a traditional feature extraction method that resolves this problem by employing an orthogonal transformation to turn a collection of potentially linear variables into a reduced set of linearly uncorrelated variables, referred to as principal components (Baburao *et al.* 2025). This

transformation allows for the depiction of high-dimensional data in a lower-dimensional space, thus enhancing the clarity of inherent data structure.

2.6.2 Machine Learning (ML)

ML is a data-centric method for pattern identification that emphasizes the use of statistical correlations among variables in a dataset to develop either generative or discriminative models.

KNN is a straightforward yet efficient ML approach that classifies a new sample by evaluating its similarity to labeled instances in the training dataset, allocating it to the most prevalent class among its k nearest neighbors. RF is an ensemble learning technique that generates numerous decision trees utilizing random subsets of samples and characteristics, subsequently amalgamating their predictions by voting or averaging to improve accuracy and resilience. SVM technique identifies an ideal linear decision border that distinguishes various classes while optimizing the margin between the boundary and the closest data points from each category (Krishnasamy *et al.* 2025).

2.6.3 Partial Least Squares-discriminant Analysis (PLS-DA)

The PLS-DA integrates the PLS technique with linear discriminant evaluation grounded in Bayesian theory to facilitate the classification of samples utilizing a diminished set. The latent variables serve as representative components that maximize the covariance between the predictor and response variables. This approach was especially appropriate for spectral data analysis, since it adeptly manages high-dimensional datasets, mitigates the likelihood of overfitting during model creation, and functions under linear modeling assumptions.

3. RESULTS AND DISCUSSION

3.1. Physicochemical Analysis of AgNPs

The synthesized AgNPs were initially characterized by UV–Vis spectroscopy. Fig. 1 illustrates that the absorption spectrum displays a prominent peak at around 528 nm, indicative of the unique surface plasmon resonance of AgNPs, thus validating their successful synthesis. In prior research, UV–Vis spectroscopy monitors oxidative stability by measuring conjugated diene (≈232 nm) and triene (≈268 nm) formation in edible oil. A lower increase in absorbance confirms the superior antioxidant performance of OG (Li *et al.* 2025).

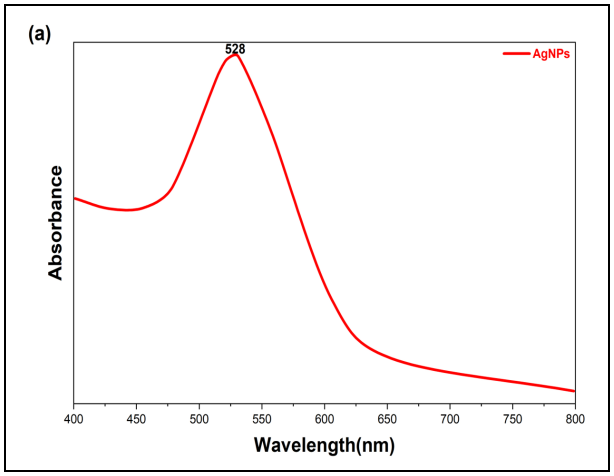


Fig. 1: UV–Vis absorption spectrum of synthesized AgNPs, showing a characteristic surface plasmon resonance peak at approximately 528 nm, confirming their successful formation

To assess the SERS efficacy of the AgNP substrate, MB was used as the probe molecule owing to its distinct Raman characteristics. MB displays distinct Raman bands at around 622 cm^{-1}, indicative of the in-plane C–C–C vibration, whilst peaks at 722 cm^{-1} and 1122 cm^{-1} correspond to in-plane and out-of-plane C–H bending modes. Additional notable bands detected at 1322 cm^{-1} and 1522 cm^{-1} correspond to aromatic ring stretching vibrations. Fig. 2a demonstrates that SERS spectra of MB at varying concentrations, using AgNP substrate, exhibit a gradual reduction in characteristic peaks at 622 cm^{-1}, 722 cm^{-1}, and 1322 cm^{-1} as concentration decreases; nevertheless, these peaks remain discernible even at 10^{-9} M, highlighting the substrate's robust signal enhancement capability. Intermittent periodic background signals were discovered during SERS measurements in specific batches of AgNPs, which were infrequently present in MB detection alone and were ascribed to the SPR action of silver nanoparticles. The SPR frequency is determined by the size, morphology, and structural configuration of the nanoparticles, while aggregation or organized assemblies may induce intricate plasmon coupling, leading to periodic Raman characteristics. Fig. 2a illustrates that both the intensities of the unique MB peaks and the periodic background signals diminish as MB concentration decreases, presumably due to the attenuated SPR intensity. Additionally, Fig. 2b illustrates a robust exponential relationship among Raman peak intensity at 1322 cm^{-1} and MB concentration, represented by the equation, $I = 7048.96 \times \exp.(\log_{10}C/0.31) - 2.96$, exhibiting a high correlation coefficient ($R^2 = 0.9943$).

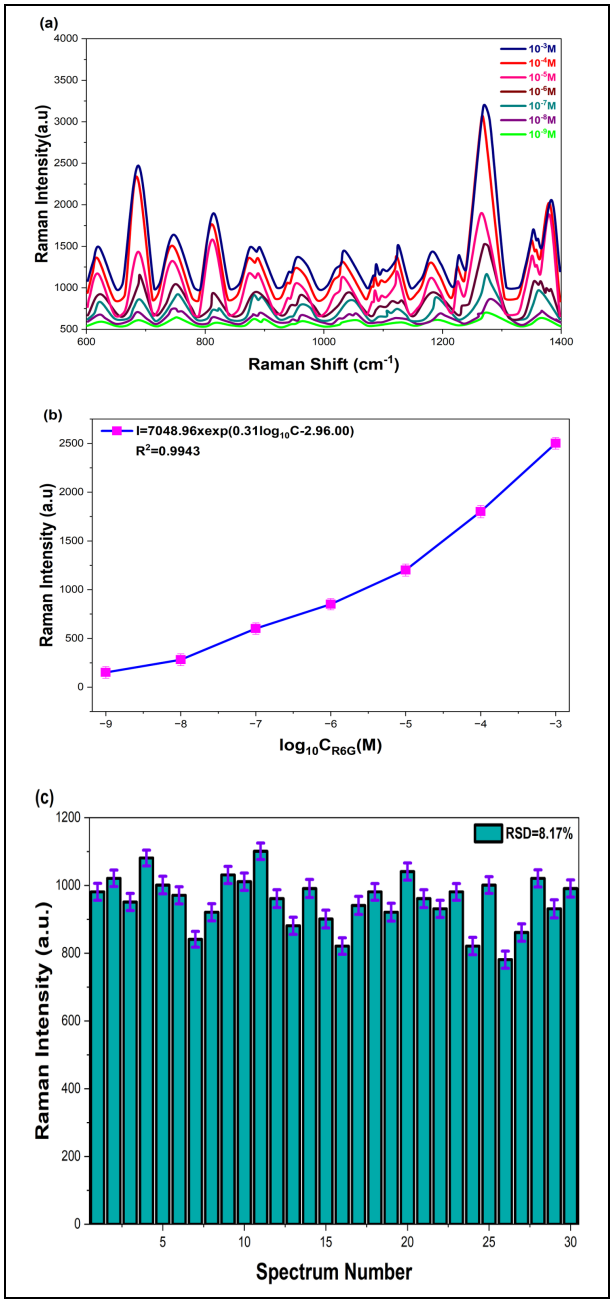


Fig. 2: (a) SERS spectra of MB at different concentrations demonstrating the signal enhancement capability of AgNP substrates, (b) exponential relationship between Raman peak intensity at 1360 cm⁻¹ and MB concentration and (c) intensity distribution across different AgNP batches indicating good substrate reproducibility

The enhancement factor (EF) of the SERS substrate was assessed by contrasting the Raman intensity of MB at a concentration of 10^{-9} M on AgNP with that of 10^{-3} M on the blank substrate. The enhancement factor was computed using a methodology documented in the literature, based on the standard formula for its determination.

$$EF = \frac{I_{SERS}}{I_{bare}} \times \frac{C_{bare}}{C_{SERS}} \quad \ldots (1)$$

In this context, I_{SERS} and I_{bare} represent the intensities of the characteristic Raman peak recorded with and without SERS, while C_{SERS} and C_{bare} denote the concentrations of MB on the AgNP-modified substrate and the blank substrate. The EF of AgNP substrate, determined from the intensity of the characteristic peak at 1360 cm^{-1}, was calculated to be 4.4×10^5. This notable improvement was due to the densely arranged silver nanoparticles on the substrate surface, which effectively enhanced the local plasmon resonance and hence increased the Raman signal intensity. As reported in previous research, Raman bands at ~1655 cm^{-1} (C=C) showed smaller intensity reduction, while peroxide-related signals (~1745 cm^{-1}) increased less in OG-treated emulsions (Delfanian *et al.* 2023).

To assess the repeatability of the substrate, five distinct AgNPs were produced under uniform experimental circumstances. Each batch served as a SERS substrate for the independent measurement of 10^{-6} M MB solution, resulting in six spectra recorded per batch and a cumulative total of 30 SERS. The spectra were statistically examined by creating a histogram of the peak intensities. Fig. 2c illustrates that the histogram of the characteristic peak at 1322 cm^{-1} has a relative standard deviation of 8.17%, signifying commendable stability and repeatability of the synthesized substrates.

The exceptional homogeneity of the substrate was additionally confirmed using SERS mapping research. Raman intensities of MB at 622, 722, and 1222 cm^{-1} were measured from 1600 sites within a randomly chosen area of 20 × 20 μm². In previous research, AgNP-based SERS has been integrated after fast-frozen solid–liquid extraction (FF-SLE) to enable the rapid detection of synthetic antioxidants in edible oils, with enhancement factors > 10^6. Quantitative SERS models can achieve linearity ($R^2 \geq 0.99$) and detection limits below 100 ng/g, comparable to HPLC-UV performance (Wang *et al.* 2023a).

3.2 SERS-based Trace Antioxidant Detection

Fig. 3 illustrates the normal Raman spectra of SM, OG, and FA powders in the range of 300–1800 cm^{-1}. In the case of SM, the most noticeable Raman bands include 605, 625, 890, 1205, and 1585 cm^{-1} that denote C-O, C-C, ring deformation, and aromatic vibrations. Peptide-like peaks can be observed at 520, 610, 900, 1210, and 1590 cm^{-1} in the case of OG, which can be ascribed to C-C peaks, phenyl group vibrations, and aromatic rings. FA shows a specific set of bands at 610, 630, 885, 1200, and 1580 cm^{-1}, and these are mainly related to C-O stretching, C-C stretching, and aromatic ring breathing vibrations. To analyze quantitatively, the strongest and most resolved characteristic peaks were selected, namely those at approximately 605 cm^{-1}, 520 cm^{-1}, and 620 cm^{-1}, which exhibit minimal spectral overlap and are therefore the most suitable marker peaks

for discrimination studies. In prior studies, Raman spectroscopy of sesamol showed a rapid reduction in 2,2-diphenyl-1-picrylhydrazyl (DPPH)-related peak intensity (~1330 cm^{-1}) within 5–10 min in methanol, indicating strong radical scavenging activity (La *et al.* 2021).

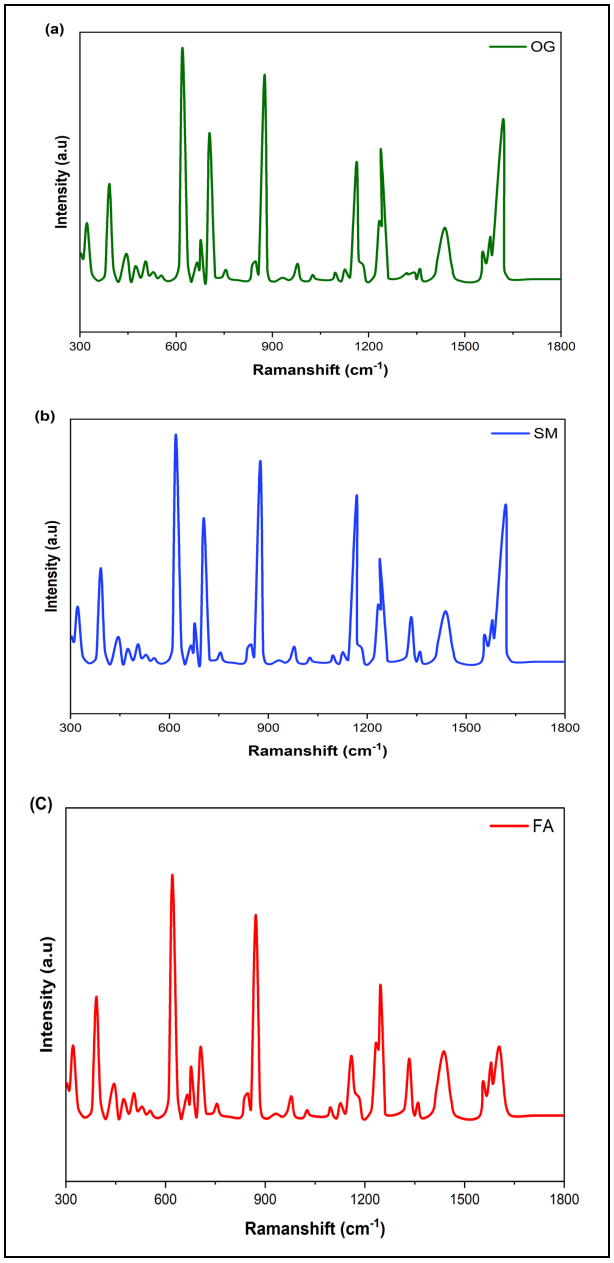


Fig. 3: Raman spectra of (a) octyl gallate (OG), (b) sesamol (SM) and (c) ferulic acid (FA), showing characteristic peaks used for the identification and differentiation of antioxidant compounds

Soybean oil is among the most extensively utilized edible oils globally and, owing to its elevated levels of unsaturated linolenic acid, is especially prone to oxidative deterioration. To examine the correlation between SERS signal intensity and antioxidant concentration, soybean oil samples infused with SM, OG, and FA were generated throughout a concentration spectrum of 400 to 50 mg/kg. Given the intricate array of food oils, all samples were subjected to a pretreatment process that included centrifugation, extraction, and subsequent amalgamation with AgNPs, before SERS measurement. This approach facilitated the acquisition of SERS at varying concentrations in oil systems. To guarantee reproducibility, multiple random regions exhibiting diverse optical contrast were chosen on the substrate using a Raman microscope for spectral acquisition. The spectrum denotes several measurement points, encapsulating overall substrate efficiency. Fluctuations in spectral intensity across several locations were predominantly noted at signal magnitude rather than in typical positions, suggesting that signals emanated from identical analyte molecules situated at distinct amplification. To mitigate potential crystallization at elevated antioxidant concentrations, Raman spectra from crystalline regions were compared with standard SERS spectra (Figs. 4–6). The notable discrepancies in peak positions affirm that the observed signals originate from enhanced Raman scattering on SERS, rather than from crystalline surfaces.

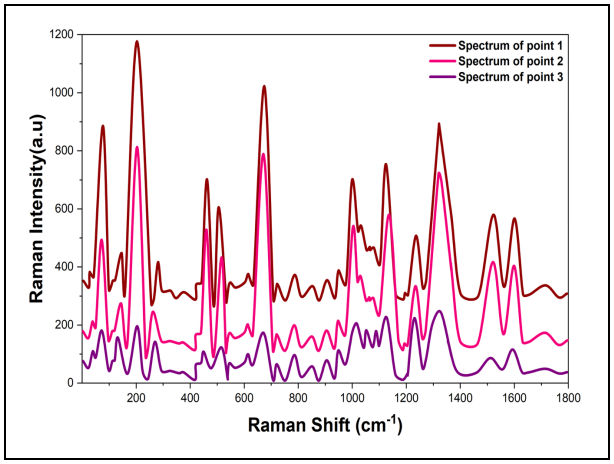


Fig. 4: Comparison of average Raman signals from different locations on the substrate containing high-concentration OG

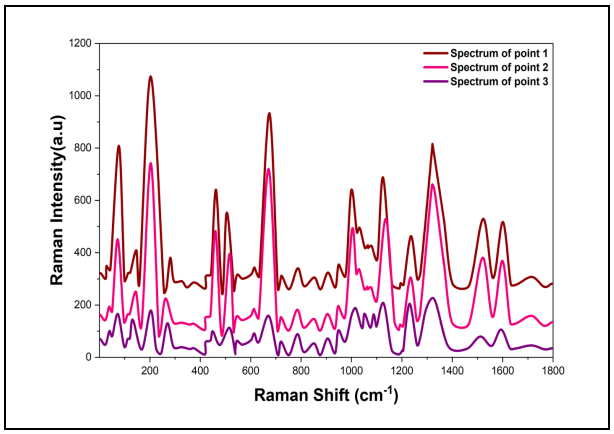


Fig. 5: Comparison of average Raman signals from different locations on the substrate containing high-concentration SM

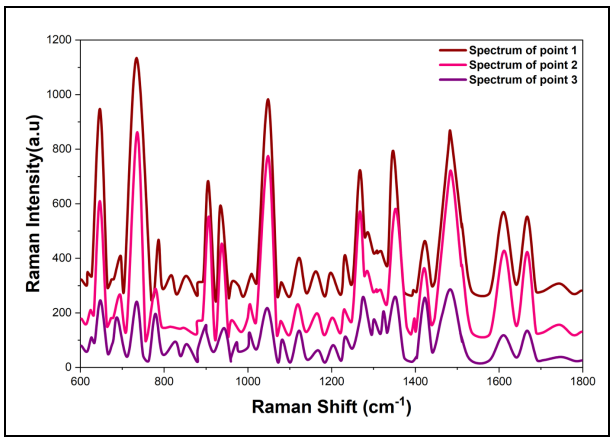


Fig. 6: Comparison of average Raman signals from different locations on the substrate containing high-concentration FA

Figs. 7(a–c) illustrate that averaged SERS at different doses distinctly exhibits the distinctive peaks of SM at 611 cm^{-1}, OG at 511 cm^{-1}, and FA at 611 cm^{-1}, which remain resolvable even at 20 mg/kg, thereby underscoring the substantial enhancing potential of AgNPs. Furthermore, the SERS intensities of these distinctive peaks demonstrate robust linear correlations with antioxidant content between 400 and 50 mg/kg, with R^2 values of 0.9383, 0.9761, and 0.9607, as illustrated in Figs. 7(d–f).

$$Limit\ of\ Detection\ (LOD) = \frac{3\sigma}{S} \quad \ldots (2)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - x_{\text{average}}\right)^2} \quad \ldots (3)$$

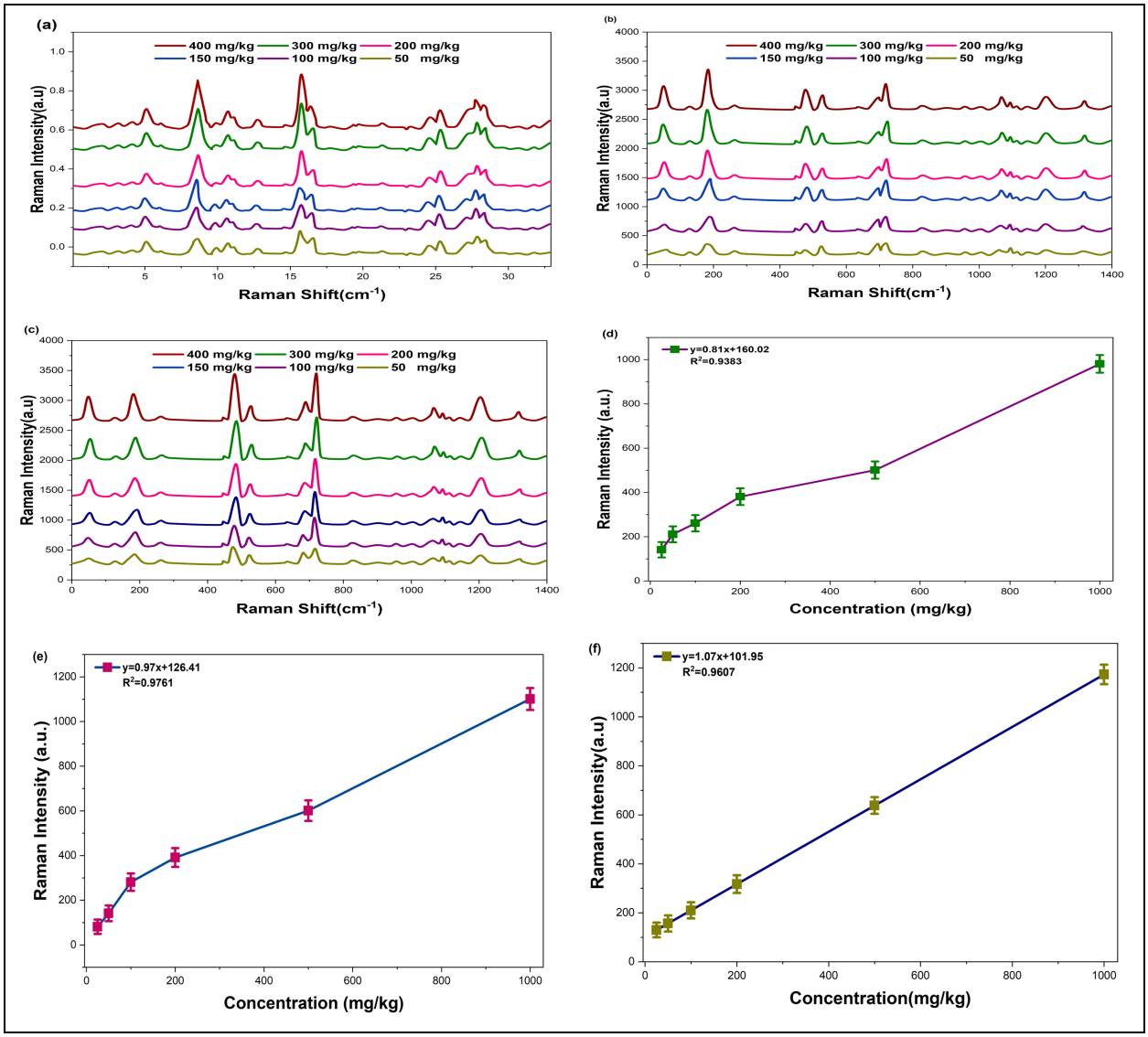


Fig. 7: Concentration-dependent SERS spectra of (a) SM, (b) OG, and (c) FA, with corresponding linear calibration curves of peak intensity vs. concentration for (d) SM, (e) OG and (f) FA

In Eq. 2, S is the slope of linear regression, whereas σ signifies the standard deviation of the blank. A detection threshold of 3 antioxidants was determined as follows: 9.711 mg/kg for SM, 8.107 mg/kg for OG, and 7.345 mg/kg for FA.

The distinct characteristics of antioxidants such as OG, FA, and SM influence interactions with silver nanoparticles, resulting in variances in SERS. OG possesses a stable structure, facilitating homogeneous adsorption onto the nanoparticle surface, which leads to reduced error bars and a low RSD value. FA and SM exhibit distinct interactions with silver nanoparticles, resulting in increased error margins and greater signal fluctuations at varied doses.

3.2.1 Comparative Spectral Features of Antioxidants

Prior experimental findings showed that the SERS approach was proficient at identifying individual antioxidants in consumable oils. Nonetheless, owing to the synergistic interactions among several antioxidant pathways, the incorporation of antioxidants into oils was not restricted to a single type. Consequently, they are essential to evaluate the feasibility of accurately differentiating among these various addition procedures.

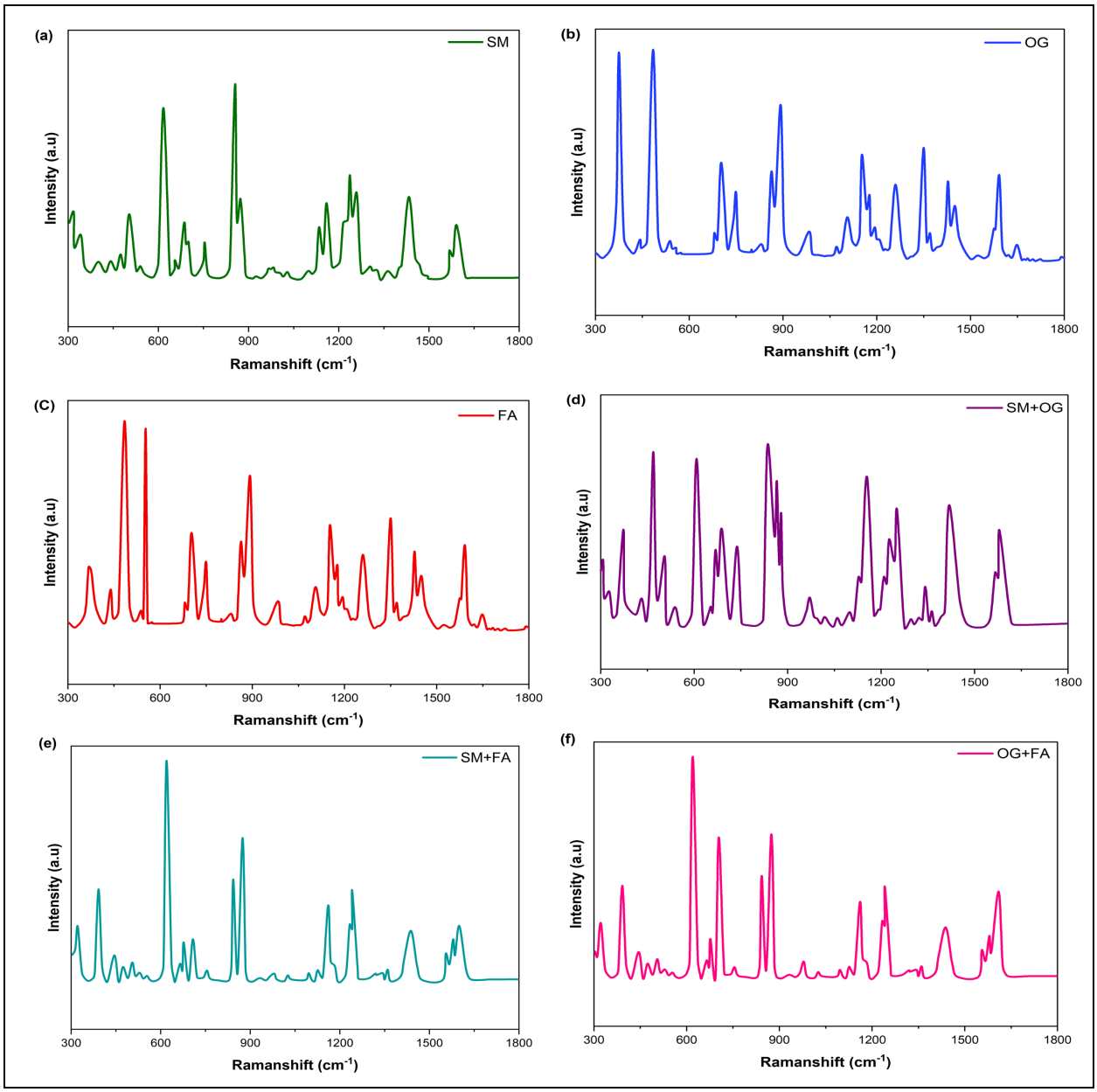


Fig. 8: Raman spectra of (a) SM standard, (b) OG standard, (c) FA standard, (d) SM–OG mixture, (e) SM–FA mixture and (f) OG–FA mixture

Characteristic peaks for the SM standard solution were observed at 500, 618, 855, 1238, and 1436 cm^{-1}. Characteristic peaks for OG were detected at 376, 486, 893, 855, and 1347 cm^{-1}. For FA, significant peaks were seen at 411, 533, 544, 655, and 955 cm^{-1}. The outcomes closely correspond to the spectra of powdered materials. At combined OG and SM, peaks at 470, 605, and 1251 cm^{-1} were identified, matching OG, whereas peaks at 834 and 1150 cm^{-1} were attributed to SM. At mixed solutions of FA and SM, as well as FA and OG, distinct peaks of all antioxidants were seen (Fig. 8).

PCAs were conducted on the obtained Raman spectra of solutions to analyze the data, resulting in a three-dimensional component score plot that illustrates the distinctions across various antioxidants (Tian *et al.* 2025). Fig. 9 shows the three-dimensional PCA score plot of the entire spectral data. The initial three major factors explained cumulative variance of 95.19, which was contributed by 65.69, 26.62, and 2.88% by PC1, PC2, and PC3, respectively. The score distribution shows that the six antioxidants are clustered around discrete and tight clusters, between which there is no inter-class overlap, and there is a clear separation between the classes. The dominance of PC1 shows that the main spectral variability of the antioxidants is well captured on this axis, and PC2 can further increase the discrimination of closely related samples. PC3 is the means that introduces another layer of variation in three-dimensional space that enhances the ability to see minor differences. The observable clear separation in the plot indicates that the spectral characteristics in the chosen range have enough discriminating information to distinguish all six antioxidants with a high level of accuracy, and thus, shows that the PCA-based classification method is robust and effective.

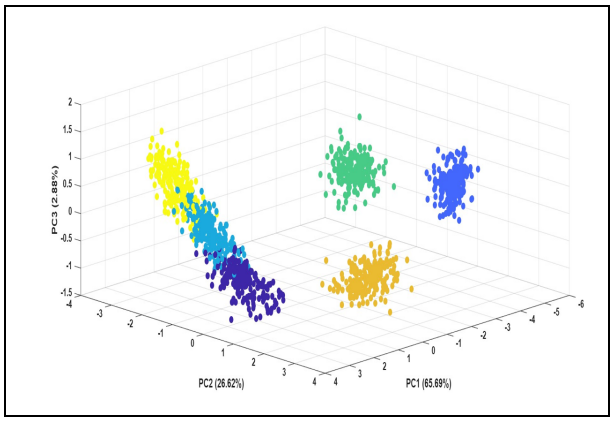


Fig. 9: PCA of the full Raman spectrum, showing a 3D PCA plot of six antioxidants

3.3 Classification and Identification Model Development

Six distinct antioxidant-infused oil solutions were developed to categorize types of antioxidants included in edible oil, hence creating a classification model dataset. Fig. 10 displays raw SERS spectra acquired at oil solutions containing individual antioxidants at different doses. Variability exists in the SERS spectra recorded for various locations for identical concentration, particularly for FA and SM. This variance results from the influence of the oil's environmental context and complex matrix.

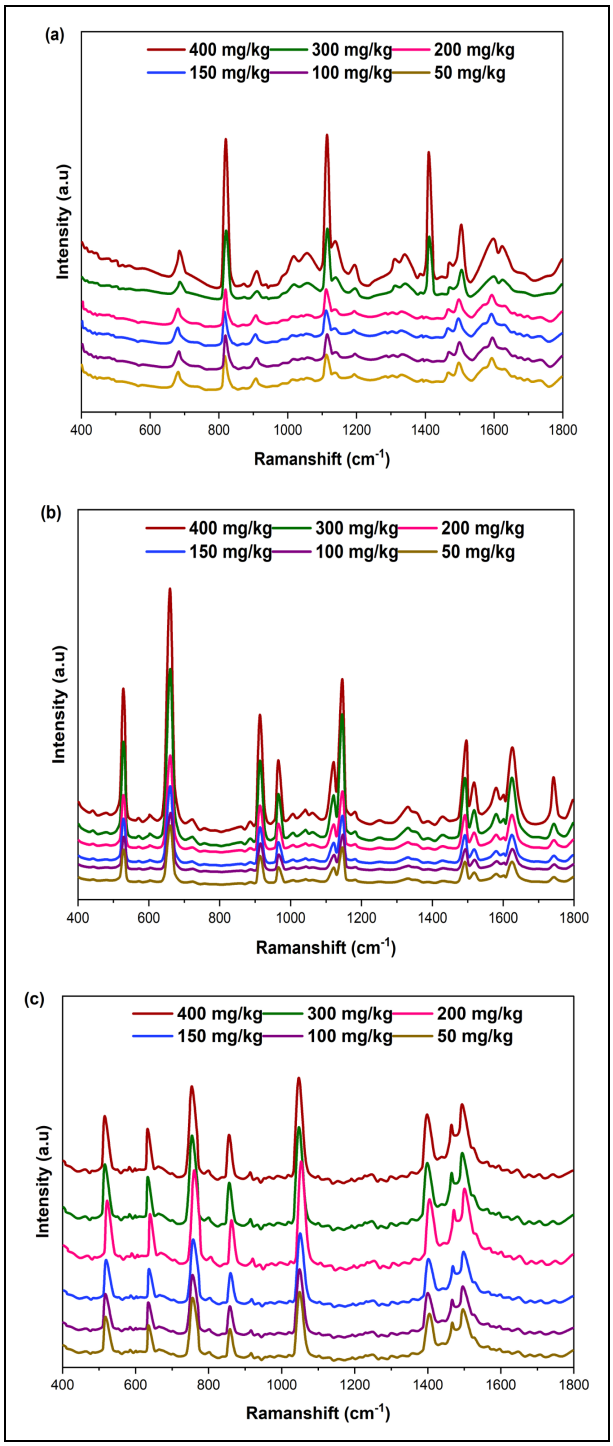


Fig. 10: SERS of (a) SM, (b) OG and (c) FA at various concentrations

Fig. 11 displays SERS findings for binary antioxidants. At identical concentrations, the percentage variance among various spectra is minimal, signifying substantial stability and consistency of spectra. Nonetheless, the overlap of characteristic peaks from two antioxidants leads to minimal variation in the peak intensities of certain features with concentration, thereby introducing uncertainty in the spectral data. To optimize the use of spectral data for constructing a classification model, preprocessing activities are essential. The preprocessing procedures encompass smoothing to refine the curves and minimize noise, and normalization to standardize spectral data to a consistent scale.

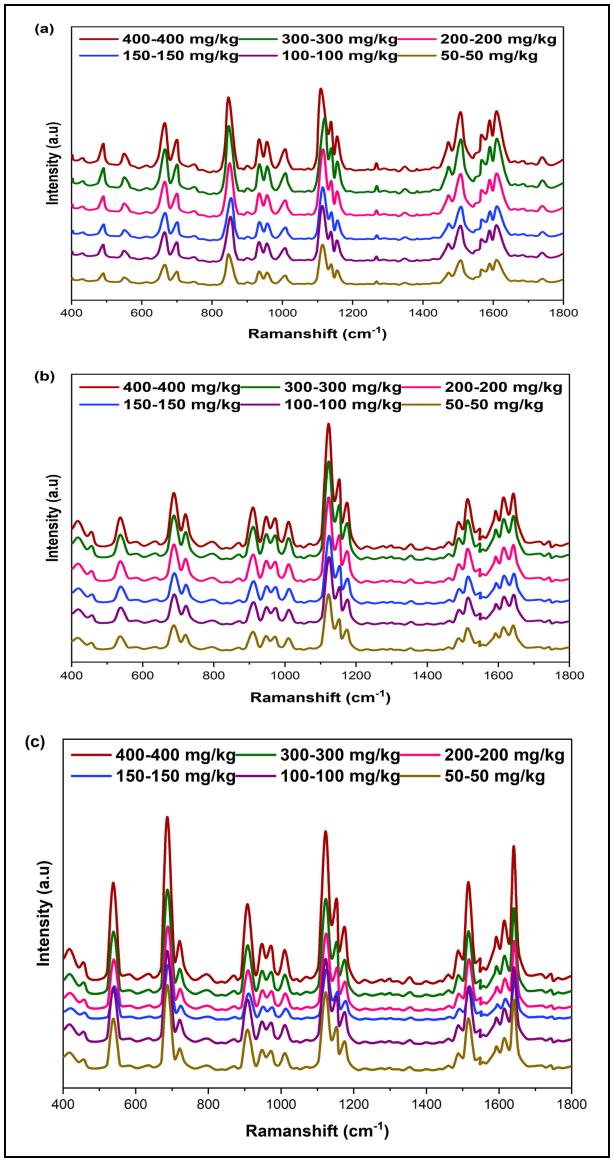


Fig. 11: SERS of (a) SM + OG, (b) SM + FA and (c) OG + FA at various concentrations

To enable the differentiation of response characteristics of individual components, the average SERS spectra at various concentrations were analyzed for FA, OG, SM, and binary mixes. The SERS intensity of composite systems and individual antioxidants shows considerable variation with concentration. This offers critical information for future development of categorization methods.

All SERS spectra used for model development were collected from antioxidant-contaminated soybean oil samples with concentrations ranging from 50–400 mg/kg. The dataset was divided into training (80%) and testing (20%) subsets to evaluate the classification performance of machine learning models.

Table 1. Dataset preparation for machine learning classification of antioxidants in soybean oil using SERS spectral data

S. No.	Sample Size	Antioxidants	Test Set	Training Set
1	180	SM	36	144
2	180	OG	36	144
3	180	FA	36	144
4	180	SM + OG	36	144
5	180	SM + FA	36	144
6	180	OG + FA	36	144

RF, SVM, and KNN classification models were initially developed. KNN is a straightforward and efficient pattern recognition technique. A primary concern with the model was the determination of the k-value.

The efficacy of random forest was contingent upon two pivotal parameters: n_estimators and max_depth. The ideal parameter configuration was determined by a grid search approach, with n_estimators established at 40. A grid search approach was utilized on the SVM model, revealing the ideal parameter combination of gamma = 0.167 and C = 100 for the SVM algorithm's key parameters.

The three classifications, RF, SVM, and KNN, were developed utilizing SERS spectra. Matrices were generated to illustrate the classification outcomes of several models at the test set. Classification errors were noted for each category of antioxidant in the KNN model. The SVM and RF models produced superior classification outcomes. The RF erroneously identified 6 OG samples as SM, 1 SM + FA sample as SM + OG, and 4 SM as OG, correctly recognizing other samples.

$$L\,(y, y_{\text{real}}) = -\frac{1}{M}\sum_{i=1}^{M} y_{\text{real_i}} \times \log\left(softmax\,(y_i)\right) \quad \ldots (4)$$

$$\text{softmax}\,(y_i) = \frac{e^{y_i}}{\sum_{i=1}^{M} e^{y_i}} \quad \ldots (5)$$

$$\text{Accuracy} = \frac{\sum_{m=1}^{M} n_{mm}}{N} \times 100\% \quad \ldots (6)$$

In the formula, n_{mm} signifies the quantity of samples classified properly in class *m*. *M* indicates the total number of classes, and *N* specifies the total number of samples.

Table 2. Performance comparison of machine learning models for antioxidant classification based on SERS spectral data

	SVM	RF	KNN
Accuracy	98.24%	95.63%	89.41%
Precision	98.30%	96.02%	89.90%
F1-score	98.18%	95.74%	89.12%
Recall	98.05%	95.48%	88.96%

Among the evaluated models, SVM achieved the highest classification accuracy (98.24%), demonstrating its strong capability to handle high-dimensional SERS spectral data and effectively separate antioxidant classes through optimal decision boundaries. The RF model obtained an accuracy of 95.63%, benefiting from ensemble learning that reduces overfitting and improves prediction stability. In contrast, KNN showed lower accuracy (89.41%), which can be attributed to its sensitivity to spectral noise and dependence on distance-based similarity measures. Despite the high overall performance, minor misclassifications were observed in certain mixed antioxidant samples due to overlapping spectral features and variations in signal intensity across the substrate. These results indicate that while the SERS–machine learning framework provides reliable classification performance, further improvements, such as larger datasets, advanced spectral preprocessing, and model optimization, may enhance robustness and predictive accuracy in complex food matrices.

3.4 Qualitative Assessment of Antioxidants in Oils

Conventional analytical techniques such as high-performance liquid chromatography (HPLC) and gas chromatography–mass spectrometry (GC–MS) are commonly used for antioxidant detection due to their high sensitivity and accuracy. However, these methods require complex sample preparation, longer analysis time, and expensive instrumentation. In contrast, the proposed SERS combined with a machine learning approach enables rapid spectral acquisition and automated classification, making it suitable for fast screening and real-time monitoring of antioxidant levels in edible oils. Therefore, the integration of SERS spectral analysis with machine learning provides an efficient alternative for food safety and quality monitoring.

3.4.1 Development of Analytical Model Dataset

The 50, 100, and 150 mg/kg for individual antioxidants, such as OG, FA, and SM, were regarded as acceptable concentrations compared to 200, 300, and 400 mg/kg. After categorizing all spectrum data as excessive or within allowed limits, the data were randomly mixed and partitioned in an 80:20 ratio.

3.4.2. Qualitative Analysis Model Development

The use of the PLS-DA score plot shows clear and well-defined clustering of the four sample groups, that is, FA, OG, OG + FA, and OG + OG. Latent Variable 1 (LV1) is the primary contributor to inter-class discrimination in the horizontal axis, whereas Latent Variable 2 (LV2) supports further separation in the vertical axis.

The samples of FA are clearly concentrated on the negative side of LV1, which shows a high level of compositional homogenization and a lack of overlap with other samples. The distribution of OG samples is positive and separated on LV2 with a moderate spectral variance and identity of classes. The OG + FA cluster is a very small cluster located in the upper positive quadrant, with very great synergistic spectral characteristics compared to those of individual components. In the meantime, OG + OG samples are evidently segregated on the far positive LV1 region, making the samples highly concentrated and isolated, which demonstrates the presence of strong discriminative behavior.

There is little overlap among clusters, which means that the model is highly robust and that the latent variables are well extractable. The reduced size within each of the classes further implies that SERS spectral signatures are highly reproducible. In general, the PLS-DA model is effective in the achievement of multi-class discrimination with a good inter-group separation and high classification reliability.

Using a decision threshold of 200 mg/kg, SVM modeling was applied to the pre-processed spectral dataset in order to categorize antioxidant samples. Samples were deemed defective if their antioxidant concentrations were less than 200 mg/kg, and non-defective if their concentrations were 200 mg/kg or higher. The classification performance of the SVM model is presented in Fig. 13.

The SVM results demonstrate strong discriminative capability for detecting excessive antioxidant concentrations. The identification of FA and

SM achieved 100% prediction accuracy, indicating clear separation between defect and non-defect samples. However, minor misclassifications were observed in certain binary antioxidant systems. In the case of SM + OG, two defect samples were incorrectly predicted as over-limit. For OG + FA, one defect sample was misclassified as non-defect, and one non-defect sample was predicted as defect. Despite these limited errors, the overall classification accuracy remained high.

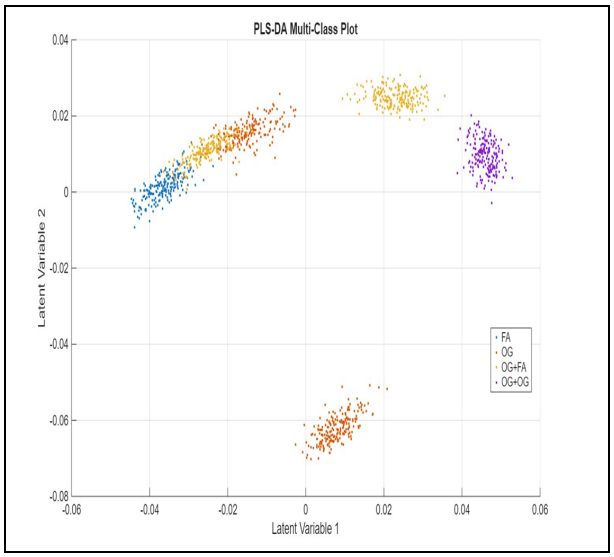


Fig. 12: PLS-DA score plot showing the classification of antioxidant samples (FA, OG, OG + FA and OG + OG) based on LV1 and LV2

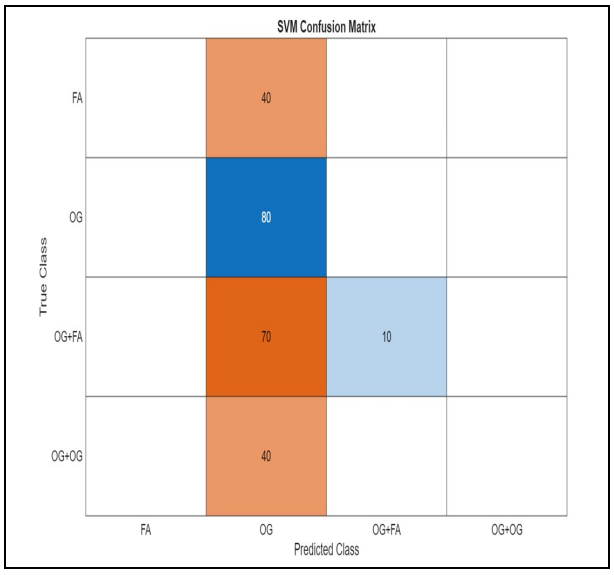


Fig. 13: (a) PLS-DA detection of excess levels of SM, OG, FA and their binary mixtures

These findings confirm that the combination of SERS spectral features and the SVM classification model provides an effective and reliable approach for identifying excessive levels of multiple antioxidant systems.

4. CONCLUSION

Unregulated and excessive fortification of edible oils with antioxidants poses potential health risks and complicates conventional analytical detection due to complex oil matrices and time-consuming analytical procedures. In this study, silver nanoparticles (AgNPs) were synthesized and used as surface-enhanced Raman spectroscopy (SERS) substrates for the detection of sesamol (SM), octyl gallate (OG), and ferulic acid (FA) in soybean oil. The synthesized AgNPs exhibited a surface plasmon resonance peak at 528 nm, with an enhancement factor of 4.4×10^5 and good reproducibility (RSD = 8.17%). Quantitative analysis showed strong linear correlations between SERS intensity and antioxidant concentration in the range of 50–400 mg/kg, with R^2 values of 0.9383 (SM), 0.9761 (OG), and 0.9607 (FA) and limits of detection of 9.711 mg/kg (SM), 8.107 mg/kg (OG), and 7.345 mg/kg (FA). Machine learning models, including SVM, RF, and KNN, were employed for antioxidant classification, where SVM achieved the highest accuracy of 97.11%, demonstrating strong capability in handling SERS spectral data. These findings demonstrate that the integration of SERS with machine learning provides a rapid, sensitive, and reliable analytical approach for monitoring antioxidant levels in edible oils, offering significant potential for practical food safety monitoring and quality control in the edible oil industry. In future research, the proposed SERS–machine learning framework can be further developed into real-time detection platforms and portable sensing devices for on-site food safety monitoring, enabling rapid screening of antioxidant levels in industrial and regulatory environments.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

COPYRIGHT

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

REFERENCES

Ahmad, I., Younas, Z., Mashwani, Z. U. R., Raja, N. I. and Akram, A., Phytomediated Selenium Nanoparticles Improved Physio-morphological, Antioxidant, and Oil Bioactive Compounds of Sesame under Induced Biotic Stress, *ACS Omega*, 8(3), 3354–3366(2023). https://doi.org/10.1021/acsomega.2c07084

Assad, N., Laila, M. B., Iqbal, R. M. F., Manahil, L., Iqbal, J., Naeem-ul-Hassan, M., Khuram, A., Assad, Y., Khan, M. N., Hussain, S., Kaplan, A., Al-Mohaimeed, A. M. and Abid, I., Green synthesis of silver nanoparticles using Solanum lycopersicum leaves extract for highly selective detection of mercury ions and photocatalytic degradation of methylene blue, *Discover Nano*, 21(1), 14(2026). https://doi.org/10.1186/s11671-025-04424-2

Avuloglu-Yilmaz, E., Yüzbaşıoğlu, D. and Unal, F., Investigation of genotoxic effect of octyl gallate used as an antioxidant food additive in in vitro test systems, *Mutagenesis*, 38(3), 151–159(2023). https://doi.org/10.1093/mutage/gead005

Baburao, G., Devi, K. V, Balakrishnan, V., Venkatesa Prabhu, V. P., Hariprabhu, M., Selvaraju, S., Shine, S. and Khaled, J. M., Evaluation of detoxification performance of Pithecellobium dulce seed pod powder upon acid treatment: removal of reactive red 195-A dye and kinetic analysis, *Zeitschrift Fur Physikalische Chemie*, 239(6), 885–910(2025). https://doi.org/10.1515/zpch-2024-0879

Chen, J., Wang, Y., Liu, W., Zhang, F., Li, J., Yang, H. and Bi, Y., Effects of free fatty acids and peroxide on thermal loss of sesamol and formation of its transformation products in soybean oil, *LWT*, 159, 113236(2022). https://doi.org/10.1016/j.lwt.2022.113236

Delfanian, M., Sahari, M. A. and Barba, F. J., Effect of lipophilized gallic acid on the oxidative stability of omega-3 fatty acids rich soy and cow milk, *LWT*, 190(2023). https://doi.org/10.1016/j.lwt.2023.115475

Dhillon, A. K., Lo, R., Barman, S., Jayaramulu, K. and Siddhanta, S., Nanohoop anchored plasmonic surfaces for polarity-dependent ultrasensitive sensing, *Nanoscale*, 17(31), 18325–18335(2025). https://doi.org/10.1039/d5nr01767j

Han, L., Li, J., Wang, S., Cheng, W., Ma, L., Liu, G., Han, D. and Niu, L., The inhibitory effects of sesamol and sesamolin on the glycidyl esters formation during deodorization of vegetables oils, *J. Sci. Food Agri.*, 101(9), 3605–3612(2021). https://doi.org/10.1002/jsfa.10989

Han, L., Li, J., Wang, S., Cheng, W., Ma, L., Liu, G., Han, D. and Niu, L., Sesamol can inhibit the formation of glycidyl ester in deep frying palm oil, *J. Food Proce. Preserv.*, 46(2) (2022). https://doi.org/10.1111/jfpp.16236

Kamil, D., Prameeladevi, T., Ganesh, S., Narayanasamy, N., Nareshkumar, R. and Thomas, S. P., Green synthesis of silver nanoparticles by entomopathogenic fungus Beauveria bassiana and their bioefficacy against mustard aphid (Lipaphis erysimi Kalt.), *Ind. J. Exp. Bio.*, 55(8), 555–561(2017).

Karrar, E., Sheth, S., Wei, W. and Wang, X., Determination of Phenolic Compounds in Gurum (Citrulluslanatus var. Colocynthoide) Seed Oil Obtained by Different Methods Using HPLC, *Food Anal. Meth.*, 13(7), 1391–1397(2020). https://doi.org/10.1007/s12161-020-01757-9

Keramat, M. and Golmakani, M. T., A kinetic approach to microwave-assisted auto-catalytic synthesis of sesamyl butyrate and evaluating its antioxidant activity on improving canola oil thermal stability, *LWT*, 127, 109364(2020). https://doi.org/10.1016/j.lwt.2020.109364

Krishnasamy, B., Thirugapillai, P., Rajakannu, A. and Selvaraju, S., Optimization and multi-objective analysis of tensile, flexural and impact strength in nano-hybrid bio-composites reinforced with Helicteres isora, Holoptelea integrifolia fibers, and nanographene, *Revista Materia*, 30, e20240864(2025). https://doi.org/10.1590/1517-7076-RMAT-2024-0864

La, J. W., Kim, M. J. and Lee, J. H., Evaluation of solvent effects on the DPPH reactivity for determining the antioxidant activity in oil matrix, *Food Sci. Biotechnol.*, 30(3), 367–375(2021). https://doi.org/10.1007/s10068-020-00874-9

Li, A., Zhao, M. T., Yin, F., Zhang, M., Liu, H., Zhou, D. and Shahidi, F., Antioxidant effects of gallic acid alkyl esters of various chain lengths in oyster during frying process, *Int. J. Food Sci. Technol.*, 56(6), 2938–2945(2021). https://doi.org/10.1111/ijfs.14933

Li, Q., Zhu, S., Li, Y., Liu, J., Fu, M., Xue, Z. and Yu, L., Do different chain lengths and configurations of gallate affect its antioxidant activity? *Int. J. Food Sci. Technol.*, 60(1), 6(2025). https://doi.org/10.1093/ijfood/vvae006

Liu, W., Zong, B., Wang, X., Yang, G. and Yu, J., Deep eutectic solvents as switchable solvents for highly efficient liquid–liquid microextraction of phenolic antioxidant: Easily tracking the original TBHQ in edible oils, *Food Chem.*, 377, 131946(2022). https://doi.org/10.1016/j.foodchem.2021.131946

Mohanasundaram, M., Selvaraj, S., Sampath, V., Venkatesa Prabhu, V. P., Selvaraju, S., Shine, S. and Khaled, J. M., Biosorption of Pb (II) ions through nanostructured teff straw based magnetized activated biocarbon: Aspects on modeling, optimization, and kinetics, *Global Nest J.*, 26(9), (2024). https://doi.org/10.30955/GNJ.06479

Nagpal, T., Yadav, V., Khare, S. K., Siddhanta, S. and Sahu, J. K., Monitoring the lipid oxidation and fatty acid profile of oil using algorithm-assisted surface-enhanced Raman spectroscopy, *Food Chem.*, 428, 136746(2023). https://doi.org/10.1016/j.foodchem.2023.136746

Nobili, M. D., Bernhardt, D. C., Basanta, M. F. and Rojas, A. M., Sunflower (Helianthus annuus L.) Seed Hull Waste: Composition, Antioxidant Activity, and Filler Performance in Pectin-Based Film Composites, *Frontiers Nutrition*, 8, 777214(2021). https://doi.org/10.3389/fnut.2021.777214

Pei, X., Yin, F., Zhong, X., Liu, H., Song, L., Zhao, G., Wang, Y. and Zhou, D., Antioxidative Effects of Gallic Acid Esters on Walnut Oil and its Emulsion, *J. Food Sci. Biotechnol.*, 41(6), 91–97(2022). https://doi.org/10.3969/j.issn.1673-1689.2022.06.011

Rodríguez-Blázquez, S., Pedrera-Cajas, L., Gómez-Mejía, E., Vicente-Zurdo, D., Rosales-Conrado, N., Léon-González, M. E., Rodríguez-Bencomo, J. J. and Miranda, R., The Potential of Plum Seed Residue: Unraveling the Effect of Processing on Phytochemical Composition and Bioactive Properties, *Int. J. Mol. Sci.*, 25(2), 1236(2024). https://doi.org/10.3390/ijms25021236

Shen, Y., Guo, C., Lu, T., Ding, X.-Y., Zhao, M.-T., Zhang, M., Liu, H., Song, L. and Zhou, D., Effects of gallic acid alkyl esters and their combinations with other antioxidants on oxidative stability of DHA algae oil, *Food Res. Int.*, 143, 110280(2021). https://doi.org/10.1016/j.foodres.2021.110280

Tian, J., Jiao, X., Guo, J., Yu, Q., Zhang, S., Gu, G., Sivashanmugan, K. and Kong, X., Quantitative Detection of Pyrene in Edible Oil via Plasmonic TLC-SERS Combined with Machine Learning Analysis, *Biosens.*, 15(8), 477(2025). https://doi.org/10.3390/bios15080477

Toorani, M. R., Farhoosh, R., Golmakani, M. T. and Sharif, A., The effects of temperature changes on the kinetic parameters of sesame, canola and olive oils in the presence of sesamol, *Iran. Food Sci. Technol. Res. J.*, 17(1), 121–132(2021). https://doi.org/10.22067/ifstrj.v15i4.76930

Toqeer, S., Qasim, M., Abideen, Z., Gul, B., Rasheed, M. and Khan, M. A., Chemical Composition and Antioxidant Activity of Seeds of Various Halophytic Grasses, *J. Am. Oil Chem. Soc.*, 95(10), 1285–1295(2018). https://doi.org/10.1002/aocs.12099

Wang, B., Xu, X. L., Zhang, M. Y., Bu, X. M., Wang, H. L., Shi, X. Z., Xu, X. and Chen, D., A fully green sample preparation method for synthetic antioxidants determination in edible oils based on natural feather fiber-supported liquid extraction, *J. Chromatography A*, 1698, 464004(2023a). https://doi.org/10.1016/j.chroma.2023.464004

Wang, H., Wang, C., Huang, J., Liu, Y., Wu, Y., You, R., Zhang, J. H., Lu, Y. and Shen, H., Preparation of SERS substrate with 2D silver plate and nano silver sol for plasticizer detection in edible oil, *Food Chem.*, 409, 135363(2023b). https://doi.org/10.1016/j.foodchem.2022.135363

Wang, L., Chen, W., Tian, Y., Duan, X., Yuan, Y., Wang, N., Xu, C., Liu, X. and Liu, Z., Preventive Effects of Sesamol on Deep-Frying Oil-Induced Liver Metabolism Disorders by Altering Gut Microbiota and Protecting Gut Barrier Integrity, *Mol. Nutri. Food Res.*, 66(9), 2101122(2022). https://doi.org/10.1002/mnfr.202101122

Wang, W., Xiong, P., Zhang, H., Zhu, Q., Liao, C. and Jiang, G., Analysis, occurrence, toxicity and environmental health risks of synthetic phenolic antioxidants: A review, *Environ. Res.*, 201, 111531(2021). https://doi.org/10.1016/j.envres.2021.111531

Wu, X., Zhao, X., Wang, J., Zhao, X., Wang, S. and Zhang, Y., Chemically and Electromagnetically dual-enhanced COFs-Au@AgNPs SERS sensor integrated with deep learning for ultrasensitive detection of neonicotinoid pesticides, *Anal. Chimica Acta*, 1373, 344501(2025). https://doi.org/10.1016/j.aca.2025.344501

Zhang, C., Zeng, M. H., Jin, J. W., He, T. X., Li, J., Ye, T. X., Guo, Z. Y. and Chen, X., Bimetallic Ag/Cu-based flower-like nano enzyme platform for antioxidants identification and monitor of edible correlation substance, *Microchem. J.*, 218, 115268(2025). https://doi.org/10.1016/j.microc.2025.115268

Zhang, M., Fan, L., Liu, Y. and Li, J., Effects of interface generation, droplet size and antioxidant partitioning on the oxidation rate and oxidative stability of water–in–oil emulsions: A comparison of coarse emulsions and nano emulsions, *Food Hydrocolloids*, 136, 108227(2023). https://doi.org/10.1016/j.foodhyd.2022.10822