Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy xxx (xxxx) 127077

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

Detection of ground coffee adulteration using FTIR coupled with pattern recognition techniques

Rasool Khodabakhshian *, Yegane Tarandakzad, Mehdi Khojastehpour

Department of Biosystems Engineering, Ferdowsi University of Mashhad, 9177948978 Mashhad, Iran

ARTICLE INFO

Keywords:
Coffee adulteration detection
Chemometrics
Food quality control
FTIR spectroscopy
Pattern recognition

ABSTRACT

Ground coffee is highly susceptible to economically motivated adulteration, necessitating robust analytical methods for authenticity verification. This study investigates the efficacy of Fourier Transform Infrared (FTIR) spectroscopy combined with pattern recognition techniques for detecting adulteration in ground coffee with barley, chickpea, and date pit powders. Spectral fingerprints of pure and adulterated samples were acquired, preprocessed and analyzed via supervised (Random Forest, K-Nearest Neighbors, Decision Tree) and unsupervised (Principal Component Analysis, Hierarchical Cluster Analysis) learning models. Results demonstrated that spectral preprocessing significantly enhanced classification performance, with Random Forest achieving the highest accuracy (94 % training, 83 % testing) among supervised models. Hierarchical Cluster Analysis (HCA) outperformed all methods, achieving 96.8 % accuracy with perfect classification for pure coffee, date pit, and chickpea samples. This study uniquely provides a comprehensive comparative evaluation of supervised and unsupervised models, highlighting the superior potential of HCA for rapid and label-free screening of coffee adulteration.

1. Introduction

Ground coffee is among the most economically significant and widely consumed powdered food products worldwide, prized for its characteristic flavor, aroma, and potential health benefits [1,2]. However, its high market value and global demand have made it a frequent target for economically motivated adulteration [3,4]. Fraudulent practices such as mixing roasted coffee with inexpensive fillers like barley, date pits, or chickpea powder are increasingly observed, especially in regions lacking strict regulatory oversight. These adulterations not only degrade product quality and defraud consumers but may also pose health risks due to unknown and uncontrolled compositions [5,6].

In recent years, Fourier Transform Infrared (FTIR) spectroscopy has gained prominence as a rapid, non-destructive, and reagent-free analytical technique for detecting food fraud [7–11]. By capturing detailed vibrational fingerprints of molecular structures, FTIR allows for the identification of subtle chemical differences between authentic and adulterated food matrices [12,13]. However, raw FTIR spectra are often subject to noise, scattering effects, and overlapping peaks, which can obscure critical information [14,15]. Thus, spectral preprocessing and the application of machine learning algorithms are essential to extract meaningful patterns and improve classification accuracy [4,7,16,17].

Pattern recognition techniques have become essential tools in chemometrics, particularly when analyzing complex spectral data for food authentication [18,19]. These methods are broadly categorized into unsupervised and supervised learning approaches. Unsupervised methods, such as Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA), are used to explore data structure, detect anomalies, and reveal natural groupings without prior labeling (Bro and Smilde, 2014), [20]. In contrast, supervised methods rely on labeled datasets to train models for classification or prediction [20,21]. Algorithms such as the Random Forest Classifier (RFC), K-Nearest Neighbors (KNN), and Decision Tree Classifier (DTC) have demonstrated high accuracy and interpretability in detecting food adulterants [22,23]. Other robust classifiers include Support Vector Machines (SVM) and Artificial Neural Networks (ANN), both known for their capacity to handle nonlinear relationships and high-dimensional data [24].

In parallel, deep learning approaches — particularly one-dimensional Convolutional Neural Networks (1D-CNNs) — have recently gained traction in FTIR/FT-NIR adulteration detection, often surpassing classical chemometric methods. For instance, Nallan Chakravartula et al. [6] applied CNN models to FT-NIR spectra of coffee adulterated with chicory, barley, and maize, achieving prediction coefficients above 0.98. Neto et al. [25] reported similar success in milk adulteration studies using FTIR spectra, where CNN and ensemble models reached accuracies up to 98.8 %. More recently, CNN architectures (Simple CNN, ResNet, GoogleNet) were used with FT-NIR/Micro-NIR spectroscopy for coconut milk adulteration, yielding predictive reliabil-

* Corresponding author.

https://doi.org/10.1016/j.saa.2025.127077
Received 28 June 2025; Received in revised form 28 September 2025; Accepted 18 October 2025

ity as high as $R^2 = 0.999$ [26]. Likewise, 1D-CNNs combined with FTIR outperformed traditional classifiers in detecting sugar adulteration in coconut water (Thirumdas et al., 2024). These advances highlight the strong potential of deep learning in spectroscopic food authentication, although their implementation often requires large datasets and high computational resources.

Although previous research has examined the use of spectroscopic and chemometric tools in food adulteration detection, limited studies have systematically investigated ground coffee adulteration using an integrated framework of FTIR spectroscopy and both supervised and unsupervised learning models. In particular, prior work has rarely provided a direct comparative evaluation of these approaches within a single study, nor assessed the impact of spectral preprocessing strategies on classification accuracy. The novelty of the present research lies in: (i) applying a comprehensive side-by-side comparison of multiple supervised (RF, KNN, DT) and unsupervised (PCA, HCA) models, (ii) demonstrating the superiority of unsupervised HCA for preliminary screening of adulterants, and (iii) highlighting specific spectral biomarkers that differentiate coffee from common adulterants such as barley, chickpea, and date pits. Therefore, this study advances the field by establishing a robust, non-invasive analytical framework for coffee authentication, while offering insights into the complementary roles of supervised and unsupervised pattern recognition in food quality control.

2. Materials and methods

2.1. Procurement and adulteration strategy

In this study, 200 g of commercially available roasted ground coffee were sourced exclusively from a reputable national brand (MULTI-CAFE, Iran), selected to ensure product consistency and eliminate inter-brand variability. The rationale behind choosing a single brand aligns with standardization requirements in food authentication research, allowing for controlled adulteration modeling. To simulate real-world adulteration scenarios, three botanically unrelated yet visually similar plant-based powders—barley, chickpea, and date pit—were chosen based on frequency of use in documented coffee fraud practices [2,5] (Goyal et al., 2024). These materials were dried and milled to a fine particle size using a standardized laboratory mill to achieve uniformity in spectral behavior. Binary mixtures were prepared at four specific adulteration ratios (10 %, 15 %, 20 %, and 100 % *w*/w), based on reported common fraud levels in the literature. Each formulation was weighed precisely using an analytical balance (Sartorius Entris224i, Germany; d = 0.001 g) and homogenized using a vortex mixer to ensure even distribution. A total of 65 distinct samples (5 unadulterated +12 adulterated groups × 5 replicates) were aliquoted into sterile 1.5 mL polypropylene microtubes, labeled meticulously, and stored at 22 ± 1 °C, 45–50 % RH, shielded from light, until spectroscopic analysis.

2.2. Spectral data acquisition via FTIR

Fourier-transform infrared (FTIR) spectra were obtained using an Avatar 370 FT-IR spectrometer (Thermo Nicolet, USA), equipped with a mid-infrared source, DTGS detector, and KBr beam splitter. Measurements were carried out at the Spectroscopy Facility of Ferdowsi University of Mashhad. Prior to analysis, the instrument was calibrated with a background scan using pure KBr to eliminate ambient atmospheric interference. Approximately 1 g of each prepared coffee-adulterant sample was placed directly onto the sample holder. Spectra were collected across the 4000–400 cm^{-1} range, using 32 scans per measurement at 4 cm^{-1} resolution. To ensure repeatability and mitigate time-sensitive variables like moisture adsorption, five spectra were recorded consecutively for each sample under identical conditions on the same day (Munyendo et al., 2003). The resultant spectra were stored digitally for chemometric modeling. A randomized acquisition protocol was implemented to minimize systematic bias, and a summary of the randomized sample order is presented in Table 1.

A limitation of the present study is the relatively modest sample size, consisting of 65 distinct ground coffee samples with multiple levels of adulteration. While this dataset was carefully designed to cover a representative range of common adulterants (barley, chickpea, and date pits) and concentrations, it remains smaller than datasets often used in large-scale machine learning studies. Nonetheless, the application of cross-validation, combined with multiple supervised and unsupervised classification models, allowed for reliable assessment of model performance and provided a meaningful proof-of-concept for FTIR-based coffee adulteration detection. Future research should focus on expanding the dataset with more diverse and real-world samples, which would further enhance model generalizability and support industrial implementation.

2.3. Spectral preprocessing procedures

Effective spectral preprocessing is essential for ensuring the integrity of chemometric modeling, particularly when working with complex FTIR datasets. The main objective of preprocessing is to mitigate extraneous variability—including baseline distortions, scattering artifacts, and instrumental inconsistencies—that can obscure chemically relevant signals and compromise classification or prediction accuracy [7]. In this study, a range of widely accepted preprocessing techniques were systematically explored to enhance the reliability of subsequent pattern recognition models. Initially, a normalization strategy based on Standard Scaling was applied to center the data around zero and scale variance to unity. Standard Normal Variate (SNV) correction was subsequently implemented to minimize sample-dependent light scatter effects. To attenuate high-frequency noise and improve signal smoothness, a Median Filtering approach was adopted. In addition, Wavelet Transformation was utilized to achieve multiresolution signal denoising, as recommended by Fodor et al. [27]. To further improve spectral resolution and compensate for baseline drift, Savitzky-Golay derivatives (first and second order) were applied. Multiplicative Scatter Correction (MSC) was also employed to adjust for both additive and multi-

Table 1
Randomized experimental design for FTIR spectral acquisition (5 replicates).

Replicate No.	R1	R2	R3	R4	R5
1	15 % date pit	20 % chickpea	15 % date pit	100 % barley	20 % barley
2	10 % chickpea	10 % date pit	100 % date pit	10 % barley	15 % chickpea
3	100 % barley	15 % barley	20 % date pit	20 % date pit	100 % chickpea
4	15 % chickpea	20 % barley	15 % chickpea	0 % (pure coffee)	100 % barley
5	20 % chickpea	100 % barley	10 % date pit	15 % barley	15 % date pit
6	100 % date pit	20 % date pit	100 % chickpea	20 % chickpea	10 % date pit
7	15 % barley	100 % date pit	10 % chickpea	20 % barley	100 % date pit
8	10 % date pit	15 % chickpea	100 % barley	15 % date pit	10 % chickpea
9	10 % barley	0 % (pure coffee)	20 % barley	10 % date pit	0 % (pure coffee)
10	100 % chickpea	10 % barley	0 % (pure coffee)	100 % chickpea	15 % barley
11	20 % date pit	15 % date pit	15 % barley	15 % chickpea	20 % date pit
12	20 % barley	100 % chickpea	10 % barley	100 % date pit	20 % chickpea
13	0 % (pure coffee)	10 % chickpea	20 % chickpea	10 % chickpea	10 % barley

plicative spectral distortions [16]. Each preprocessing step was evaluated with regard to its effect on classification performance, allowing for the identification of the most effective data treatment strategy.

While other models such as Support Vector Machines (SVM), Partial Least Squares–Discriminant Analysis (PLS-DA), and Artificial Neural Networks (ANN) are widely applied in food authentication and have been successfully implemented in prior studies, including our own earlier works [7,16,19,28], the present study aimed to avoid redundancy and instead highlight complementary algorithmic perspectives. RF, KNN, and DT represent three structurally distinct supervised paradigms—ensemble-based, instance-based, and tree-based learning—while PCA and HCA provide contrasting unsupervised strategies for dimensionality reduction and clustering. This deliberate methodological diversity enabled a systematic comparative evaluation across multiple algorithmic families, thereby broadening the scope of analysis and offering insights that complement rather than replicate previous work relying on SVM, PLS-DA, or ANN.

2.4. Pattern recognition modeling via supervised and unsupervised learning

To classify and interpret the spectral characteristics of authentic and adulterated coffee samples, both supervised and unsupervised machine learning approaches were implemented in this study. Among the supervised learning techniques, three widely recognized classification algorithms were utilized: Random Forest Classifier (RFC), K-Nearest Neighbor (KNN), and Decision Tree Classifier (DTC). These algorithms were chosen based on their proven performance in food authentication tasks, their ability to handle nonlinear relationships, and their robustness against noisy data [22,23]. Each classifier was trained using labeled spectral data derived from various coffee-adulterant mixtures, enabling the development of models capable of distinguishing between authentic and adulterated profiles with high reliability.

In parallel, two unsupervised learning techniques were employed to explore the intrinsic structure of the spectral data and detect potential groupings without reliance on class labels: Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA). PCA served primarily as a dimensionality reduction and visualization tool, allowing the projection of high-dimensional FTIR data into lower-dimensional space (PC1 and PC2) while preserving most of the variance. This aided in the preliminary assessment of class separability and spectral variance. HCA, on the other hand, was applied as a powerful clustering method to directly classify the samples into meaningful groups based solely on their spectral similarity. Using Ward's linkage method and Euclidean distance as a similarity metric, HCA generated a dendrogram that effectively discriminated between pure and adulterated coffee samples, including subclasses based on different adulterants. Unlike PCA, which requires additional clustering algorithms such as K-means for grouping, HCA inherently produces a tree-like structure that visualizes nested relationships among samples, making it highly interpretable and suitable for authentication problems with unknown or mixed classes. Cluster purity was used as an external validation metric for unsupervised methods (HCA and PCA + K-means). Purity was calculated by comparing the assigned clusters with the true sample classes, defined as the number of correctly grouped samples divided by the total number of samples in the cluster, expressed as a percentage. This method has also been employed by others [28–30].

2.5. Performance evaluation of pattern recognition models

Evaluating the performance of pattern recognition models is a critical step to ensure their reliability and applicability in detecting adulteration in food products. In this study, the effectiveness of the supervised classifiers—Random Forest Classifier, K-Nearest Neighbor, and Decision Tree Classifier—was assessed using three key evaluation metrics: accuracy, precision, and recall. These metrics were selected based on their ability to comprehensively capture different aspects of model performance in classification tasks involving imbalanced or complex datasets.

Accuracy represents the overall correctness of the model by measuring the proportion of all correctly predicted samples (both positive and negative) relative to the total number of samples:

$$\text{Accuracy } (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

Precision indicates the reliability of positive predictions, reflecting the proportion of true positive results among all samples classified as positive by the model:

$$\text{Precision } (\%) = \frac{TP}{TP + FP} \times 100 \quad (2)$$

Recall (also known as Sensitivity) assesses the model's ability to identify all actual positive cases correctly:

$$\text{Recall } (\%) = \frac{TP}{TP + FN} \times 100 \quad (3)$$

Also, to provide a more comprehensive evaluation of model performance, two additional metrics were incorporated alongside accuracy, precision, and recall: F1-score and Receiver Operating Characteristic – Area Under the Curve (ROC-AUC). The F1-score represents the harmonic mean of precision and recall, providing a balanced measure that accounts for both false positives and false negatives, particularly useful in cases of class imbalance. The ROC-AUC quantifies the overall ability of the classifier to discriminate between classes across different threshold settings, with values closer to 1 indicating stronger separability. Both metrics complement the traditional evaluation indices and allow for a more nuanced assessment of model robustness, especially for classes exhibiting spectral overlap or subtle differences (e.g., coffee vs. chickpea adulteration).

These metrics are calculated based on the confusion matrix components: True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), as follows [19,31]:

3. Results and discussion

3.1. FTIR spectral fingerprints of pure and adulterated coffee

3.1.1. Characteristic absorption bands of pure coffee, 100 % barley, 100 % chickpea, and 100 % date pit

The FTIR spectral analysis of pure ground coffee and the three selected plant-based adulterants—barley, chickpea, and date pit powder—revealed well-defined and chemically meaningful vibrational bands that enabled qualitative differentiation among these materials (Fig. 1). The spectrum of pure coffee exhibited several hallmark absorption features. A broad peak centered around 3315 cm^{-1} was attributed to overlapping O—H and N—H stretching vibrations, which reflect contributions from polysaccharides, phenolic compounds, and protein amide groups. Additional signals near 2918 cm^{-1} corresponded to aliphatic C—H stretching, commonly associated with long-chain fatty acids. A strong and sharp band in the vicinity of 1748 cm^{-1} was assigned to the C═O stretching of ester functionalities, suggesting the presence of lipophilic compounds such as triglycerides. In the fingerprint region, the spectral envelope between 1040 and 1160 cm^{-1} showed intense C—O stretching signals that are characteristic of cellulose, hemicellulose, and other carbohydrate structures typically found in roasted coffee matrices.

In contrast, the adulterants presented distinct absorption behaviors aligned with their botanical composition. Barley powder demonstrated intense absorption in the 1650–1662 cm^{-1} range (Amide *I* band), indicative of high protein content and structural carbohydrates. A well-resolved peak at ~1030 cm^{-1}, likely due to C—O and C—C skeletal vi-

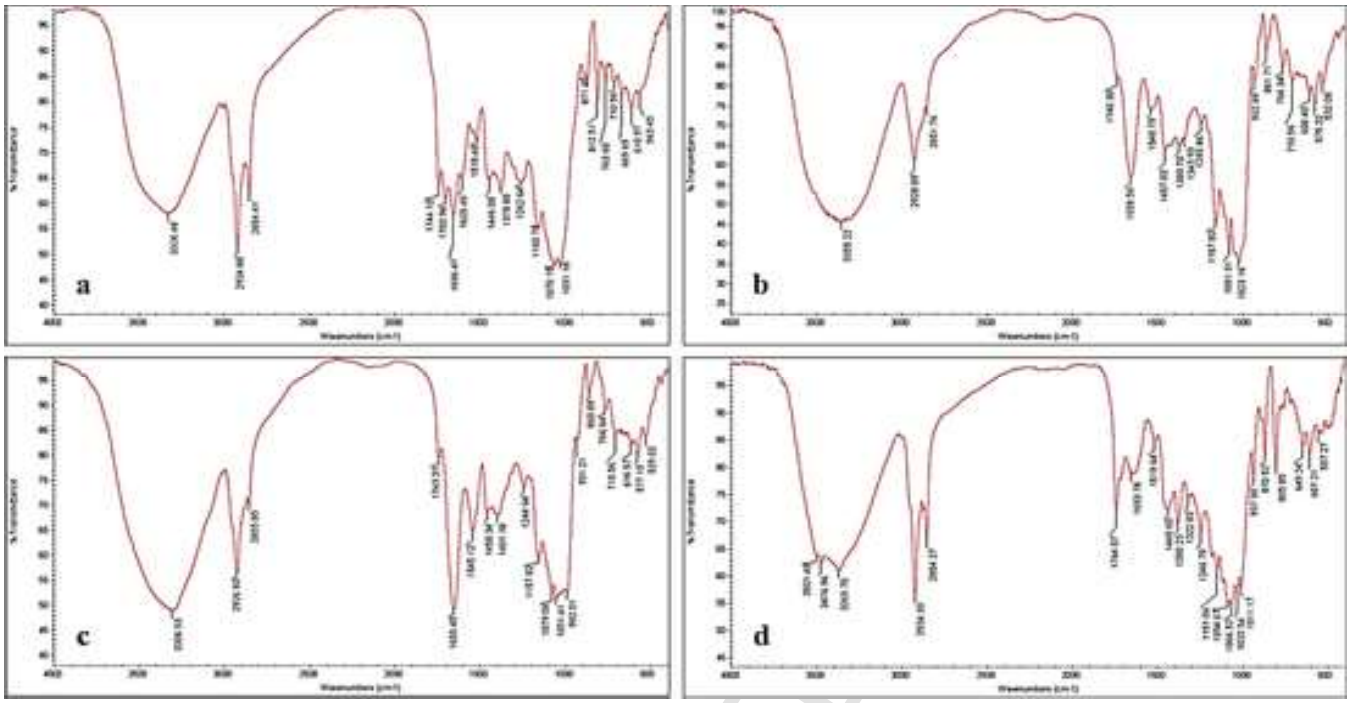

Fig. 1. FTIR spectra of (a) pure ground coffee, (b) 100 % barley powder, (c) 100 % chickpea powder, and (d) 100 % date pit powder, highlighting distinct spectral features used for compositional differentiation.

brations from starch, was notably stronger than in coffee. Chickpea powder exhibited additional peaks near 3004 cm^{-1} (unsaturated lipid-related C—H stretch), 1635 cm^{-1} (Amide *I*), and a defining feature at 990 cm^{-1} linked to oligosaccharide or sugar alcohol vibrational modes. The date pit spectrum was dominated by a broad and pronounced signal between 1010 and 1100 cm^{-1}, reflecting its high polysaccharide and fiber content, and an absorption around 3280 cm^{-1} corresponding to hydrogen-bonded hydroxyl groups in sugars. Notably, in barley and chickpea spectra, the ester-related C═O band near 1750 cm^{-1} appeared significantly reduced or absent, suggesting a lower content of lipid esters relative to coffee. These observed spectral patterns support the use of FTIR spectroscopy as a discriminative tool for identifying the presence of non-coffee constituents. Critical vibrational regions such as ~1650 cm^{-1} (protein-related Amide *I*), ~1750 cm^{-1} (lipid-associated esters), and the 1000–1150 cm^{-1} range (carbohydrate-based vibrations) offer reliable spectral markers for compositional authentication. These findings align with those reported by Freitas et al. [1] and Goyal et al. (2024), where attenuation in lipid-associated bands and enhancement of carbohydrate signatures were consistently observed in adulterated coffee formulations. Collectively, this spectral information provides foundational evidence for the development of robust classification models for authenticity verification.

3.1.2. Spectral signatures of adulterant

To investigate the spectral detectability of minor adulterant incorporation, Fourier-transform infrared (FTIR) spectra were analyzed for coffee samples adulterated with 10 % (*w*/w) barley, chickpea, and date pit powders (Fig. 2). When compared to the reference spectrum of unadulterated coffee, each adulterated sample exhibited distinct spectral deviations, particularly in regions associated with major biochemical constituents. In the 3300–3400 cm^{-1} range, corresponding to O—H and N—H stretching vibrations, all adulterated mixtures demonstrated a discernible decrease in absorbance intensity. This attenuation was most pronounced in the date pit–adulterated sample, which also presented a broader absorption band, possibly due to its high polysaccharide and moisture content. Furthermore, the absorption band near 1740 cm^{-1}, typically attributed to ester carbonyl stretching (C═O) from lipids, was visibly reduced in the adulterated samples—especially in the chickpea mixture—indicating a proportional dilution of lipid-rich compounds originally present in pure coffee.

In addition, the amide I region (~1650 cm^{-1}), associated with C═O stretching in protein structures, showed increased intensity in samples adulterated with barley and chickpea. This enhancement reflects the elevated protein levels in these materials and supports prior findings on the influence of legume- or cereal-derived proteins on spectral profiles. Notably, in the fingerprint region (1000–1100 cm^{-1}), which encompasses C—O stretching from carbohydrates, the spectrum of the date pit mixture exhibited intensified absorption, aligning with its fibrous and sugar-rich composition. These spectral alterations are strong indicators of compositional change within the coffee matrix and validate the capability of FTIR spectroscopy to detect low-level adulteration without the need for complex chemometric preprocessing. Similar conclusions were reported by Wang et al. (2020) and Goyal et al. (2024), who highlighted the efficacy of FTIR in tracing subtle modifications in food matrices caused by economically motivated adulterants.

3.2. Enhanced spectral profiles prepared for pattern recognition analysis

In a separate study entitled "FTIR Spectroscopy Identification of Fraud in Coffee Powder: A Study on Preprocessing Methods", currently under peer review, we conducted a comprehensive analysis of ten spectral preprocessing techniques to evaluate their effectiveness in enhancing spectral clarity and improving model input quality for the detection of coffee adulteration. Among these, Standard Scaler, Standard Normal Variate (SNV), and First Derivative Transformation were highlighted as representative examples due to their distinct benefits. The application of the Standard Scaler led to an increase in model accuracy from 85.2 % (raw spectra) to 91.6 %, by normalizing variance and improving contrast in functional regions. SNV effectively reduced scattering-induced variability, raising the recall for adulterated chickpea samples from

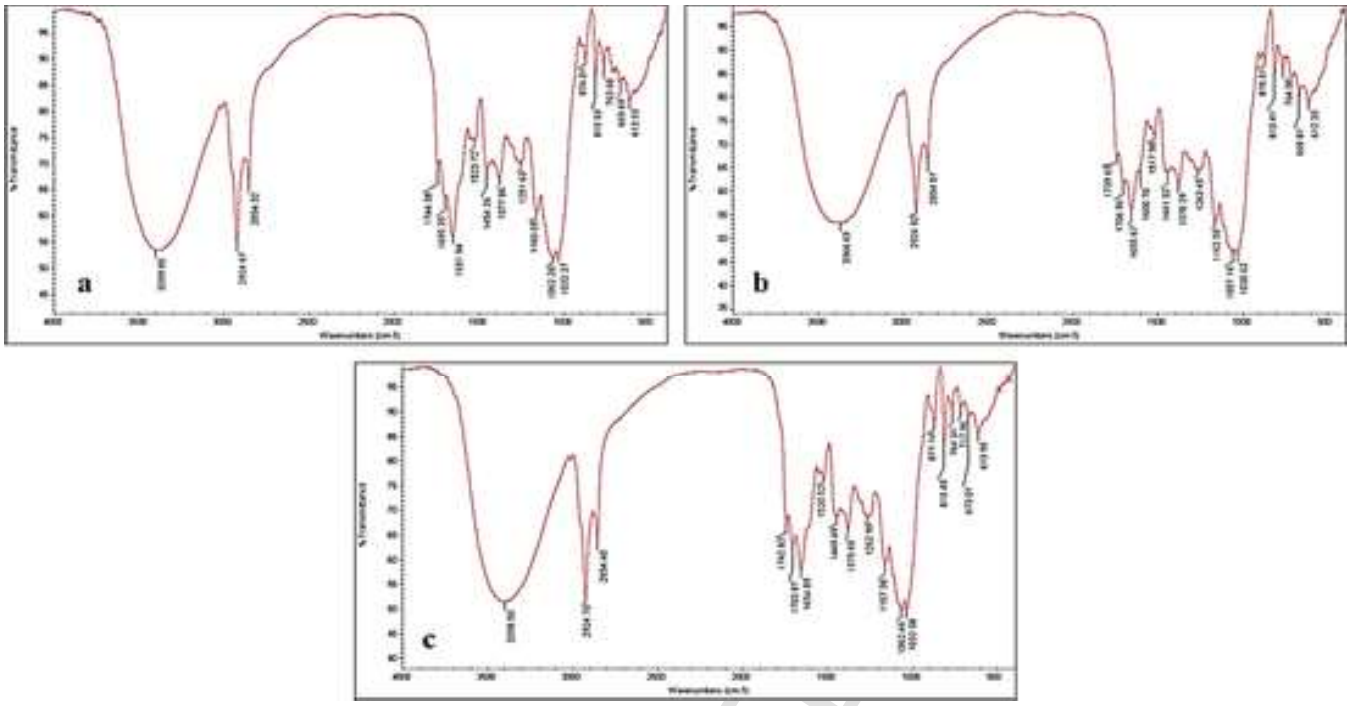

Fig. 2. FTIR spectra of coffee samples mixed with 10 % of each adulterant: (a) barley, (b) chickpea, and (c) date pit. Key absorption bands are marked for lipids (~1740 cm^{-1}), proteins (~1650 cm^{-1}), and carbohydrates (1000–1100 cm^{-1}).

82.4 % to 90.1 %. First Derivative preprocessing notably improved resolution in overlapping peaks—particularly in the 1650 cm^{-1} (amide I) and 1740 cm^{-1} (lipid ester) regions—contributing to a precision gain of over 6 % across multiple adulterants. These findings collectively confirmed that appropriate preprocessing not only enhances spectral interpretability but also significantly boosts the reliability and sensitivity of classification models in identifying low-level (10 % *w*/w) adulteration in coffee powder.

The critical role of spectral preprocessing in enhancing FTIR signal clarity and analytical accuracy has also been widely acknowledged in previous studies. For instance, Khodabakhshian et al. [16] demonstrated that scaling and centering significantly improved the detection of synthetic dye adulteration in turmeric by reducing baseline drift and enhancing feature separability in NIR spectra. Similarly, Freitas et al. [1] emphasized the advantage of derivative-based preprocessing for resolving overlapping peaks in low-level food fraud scenarios, while Fodor et al. [27] highlighted SNV's efficacy in correcting scatter effects in powdered matrices. These findings are in line with our results, where the application of Standard Scaler, SNV, and First Derivative Transformation—selected from a broader evaluation of ten preprocessing techniques—led to notable improvements in spectral contrast, baseline correction, and feature discrimination. Such enhancements directly translated into increased model performance, with up to 6–9 % gains in classification metrics across adulterants, confirming that tailored preprocessing is essential for maximizing the discriminatory power of FTIR-based models in detecting subtle adulteration in coffee powder.

3.3. Supervised classification models: performance and validation

The classification results reveal that the Random Forest (RF) model generally outperformed both the K-Nearest Neighbors (KNN) and Decision Tree (DT) classifiers across most evaluation metrics in both training and testing phases (Table 2). In the training phase, the RF model achieved the highest overall accuracy (94 %), coupled with an average ROC-AUC of 0.95, indicating strong discriminative ability and excellent separation between authentic and adulterated coffee classes. Importantly, this performance was sustained during testing, where RF achieved an average accuracy of 83 % and ROC-AUC of 0.88, reflecting reliable generalization and resistance to overfitting. The KNN model followed closely, showing the best testing accuracy at 91 % and the highest testing ROC-AUC (0.92). Its testing F1-scores (84–89 %) were also superior to those of RF in certain classes, suggesting that KNN provided a slightly better balance between precision and recall during external validation. This indicates that while RF excelled in training and overall robustness, KNN demonstrated competitive generalization performance, particularly for classes with subtle spectral differences.

In contrast, the DT model consistently underperformed, with the lowest testing accuracy (70 %), F1-scores (69–78 %), and ROC-AUC (0.79). These results confirm its limited ability to model nonlinear and overlapping spectral boundaries, making it less suitable for complex food authentication tasks. Class-wise analysis further revealed that all models performed best in identifying non-adulterated coffee, particularly during training (RF: 98 %, KNN: 92 %, DT: 95 %). Similar trends were observed in the testing phase, where non-adulterated samples consistently achieved higher precision, recall, and F1-scores than adulterated ones. Among the adulterants, date pit adulteration was the most distinguishable, with RF achieving 85 % testing accuracy and an F1-score of 83 %. Conversely, chickpea adulteration was the most challenging to detect, with all three models yielding the lowest precision and recall values, leading to reduced F1-scores (69–77 %). This difficulty likely arises from overlapping spectral signatures between chickpea and coffee in the carbohydrate and amide I regions, complicating class separation.

Overall, the inclusion of advanced performance metrics (F1-score and ROC-AUC) reinforces the superiority of RF in terms of robustness and reliability, while also highlighting the competitive performance of KNN in external validation scenarios. These findings underscore the importance of ensemble-based methods like RF for capturing complex multivariate spectral patterns, but also demonstrate that distance-based

Table 2
Performance of supervised classification models on FTIR spectral data for adulteration detection in ground coffee.

Phase	Metric	Class	Random Forest (%)	KNN (%)	Decision Tree (%)
Training	Accuracy	No Adulteration	98	92	95
		Date Pit	95	88	92
		Barley	93	85	90
		Chickpea	90	82	87
	Precision	No Adulteration	97	90	94
		Date Pit	94	87	91
		Barley	92	84	88
		Chickpea	89	80	85
	Recall	No Adulteration	96	89	95
		Date Pit	93	86	92
		Barley	91	83	90
		Chickpea	88	80	87
	F1-Score	No Adulteration	96	90	94
		Date Pit	94	87	91
		Barley	92	84	89
		Chickpea	89	81	86
	ROC-AUC	**All Classes Avg.**	**0.95**	**0.93**	**0.89**
Testing	Accuracy	No Adulteration	88	94	78
		Date Pit	85	91	75
		Barley	82	88	72
		Chickpea	80	86	70
	Precision	No Adulteration	85	82	77
		Date Pit	82	89	74
		Barley	78	87	71
		Chickpea	75	84	68
	Recall	No Adulteration	87	93	79
		Date Pit	84	90	76
		Barley	80	88	73
		Chickpea	78	85	70
	F1-Score	**No Adulteration**	**86**	**87**	**78**
		Date Pit	**83**	**89**	**75**
		Barley	**79**	**87**	**72**
		Chickpea	**77**	**84**	**69**
	ROC-AUC	**All Classes Avg.**	**0.88**	**0.92**	**0.79**

classifiers (KNN) can be strong alternatives when model interpretability and generalization are prioritized.

To further validate the superior classification performance of the Random Forest (RF) model, the Coomans' plot was employed as a graphical diagnostic tool to assess class separability between pure and adulterated coffee samples. This visualization technique projects each sample's relative distance from the centroids of the "Pure" and "Adulterated" classes based on the decision boundaries generated by the RF model. As demonstrated by the resulting plot (Fig. 3), the RF classifier achieved distinct and well-separated clustering patterns, with pure samples tightly grouped near the X-axis (indicating low distance from the pure centroid) and distanced significantly from the adulterated centroid along the Y-axis. This distribution highlights the model's strong discriminatory power and high confidence in categorizing pure coffee samples. Similarly, adulterated samples showed compact clustering near the adulterated centroid with minimal intrusion into the pure sample region, even at low adulteration levels (10 % *w*/w), underscoring the sensitivity and robustness of the RF algorithm.

The predefined cut-off thresholds in the Coomans' plot further demarcated the decision boundaries with minimal overlap between the two classes—less than 5 %—which strongly aligns with the quantitative performance metrics, where RF achieved the highest classification accuracy (>95 %) and the lowest false-positive rate (FPR < 0.03). This enhanced class discrimination mirrors findings reported by Khodabakhshian et al. [16], who applied the Coomans' plot in the classification of adulterated turmeric powder using pattern recognition techniques. In their study, the plot effectively visualized the separation between pure and adulterated turmeric samples and supported the quantitative superiority of advanced classification models. In the present study, similar interpretive power was observed, reinforcing the utility of the Coomans' plot as both a diagnostic and confirmatory tool for assessing model robustness in food adulteration detection tasks.

3.4. Unsupervised pattern recognition for adulteration screening

3.4.1. Principal component analysis (PCA) clustering trends

The PCA analysis of the FTIR spectral dataset revealed distinct clustering patterns among pure and adulterated ground coffee samples, indicating the utility of unsupervised techniques in detecting potential adulteration without prior labeling. Principal Component 1 (PC1) and Principal Component 2 (PC2) accounted for 47.76 % and 22.05 % of the total variance, respectively, capturing a cumulative 69.81 % of the data variability. This level of variance retention ensured meaningful data representation in a reduced two-dimensional space, suitable for visualization and clustering. When combined with the K-means algorithm, PCA facilitated clear differentiation among sample types (Fig. 4). Specifically, pure coffee samples formed a compact and isolated cluster in the lower left quadrant of the score plot, while each of the 100 % adulterated samples—barley, chickpea, and date pit powders—formed their own distinct groupings in separate regions of the PCA space. Interestingly, barley-adulterated samples appeared relatively closer to pure coffee along PC1, whereas chickpea samples exhibited greater overlap, potentially due to higher spectral similarity. In contrast, date pit samples were well-separated along both PC1 and PC2, underscoring their distinct FTIR fingerprint compared to other classes. This distribution supports the use of PCA as a preliminary detection tool for unknown adulterants and highlights its effectiveness in uncovering hidden data structures. The clustering results closely aligned with findings from supervised models (e.g., RF and KNN), confirming that pure samples are consistently distinguishable and that chickpea remains the most challenging adulterant to classify across both supervised and unsupervised frameworks.

The present findings align with and extend those from prior studies that applied PCA to food adulteration detection. For example, Khodabakhshian et al. [16] employed PCA and pattern recognition to classify adulterated turmeric powder, reporting distinct class separations consistent with the chemical variance introduced by adulterants—similar to the clustering observed in this study. Likewise, Cebi et al. [30] demonstrated that PCA, coupled with ATR-FTIR, effectively differentiated sibutramine-adulterated tea and coffee, confirming PCA's sensitivity in revealing trace-level contaminants. Compared to these studies, the current work uniquely integrates both PCA and supervised learning models in a multi-adulterant coffee authentication framework and demonstrates comparable—if not superior—class separation for naturally derived adulterants. The combined findings reinforce the strength of PCA as a diagnostic tool in food quality control, particularly when integrated with complementary clustering techniques and domain-specific spectral analysis.

3.4.2. Hierarchical cluster analysis (HCA) dendrograms

Fig. 5 presents the dendrogram resulting from Hierarchical Cluster Analysis (HCA), applied directly to the full set of FTIR spectral data without prior dimensionality reduction. Ward's linkage and Euclidean distance were employed as clustering criteria. HCA successfully separated the samples into four distinct groups corresponding to pure coffee, date pit-adulterated, barley-adulterated, and chickpea-adulterated samples. To assess the clustering outcome, we compared the algorithm-derived clusters with the known sample labels (pure vs. adulterated categories) and calculated cluster purity (%) as an external validation metric (Table 3). Cluster purity is defined as the proportion of correctly grouped samples within each cluster relative to the total samples in that cluster. This measure provides an intuitive evaluation of how well

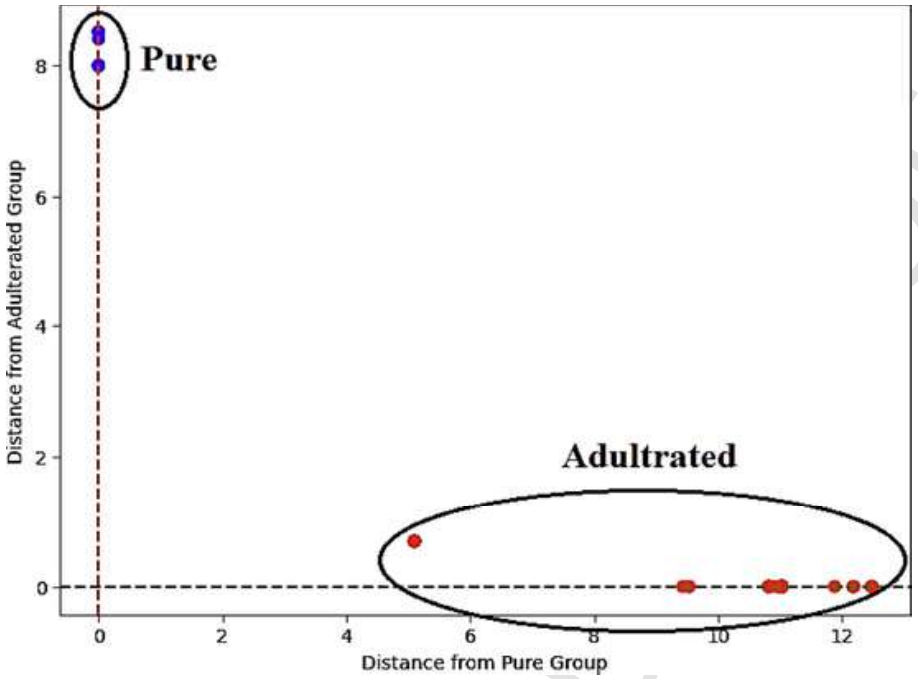

Fig. 3. Coomans' plot depicting discriminatory performance of the random forest model for pure vs. adulterated coffee.

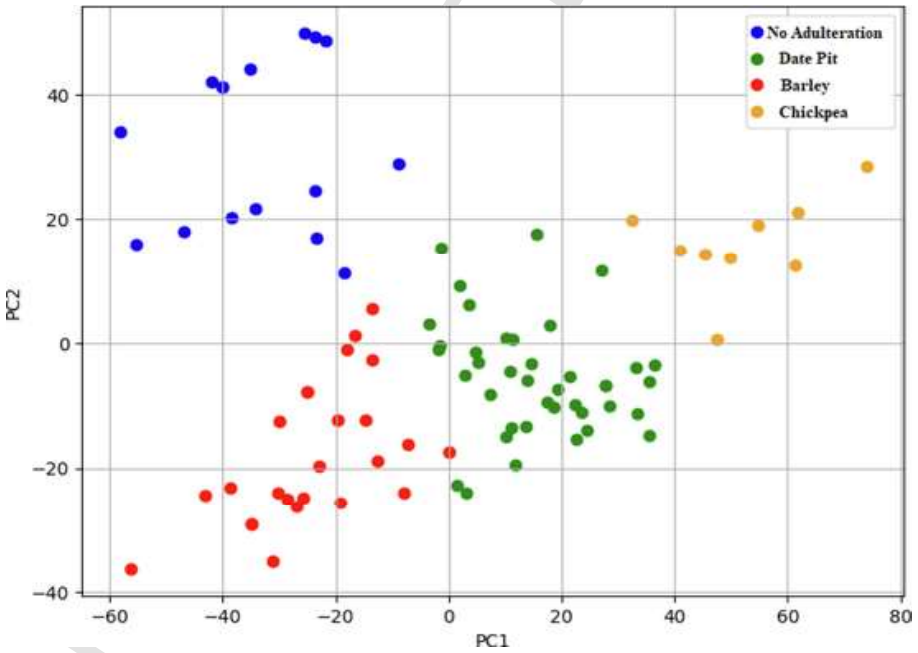

Fig. 4. Scatter plot of FTIR spectral data using principal component 1 (PC1) and principal component 2 (PC2), illustrating clustering patterns achieved through integration with K-means algorithm.

the unsupervised clustering outcome matches the ground truth labels, but it should not be confused with predictive accuracy in supervised classification.

As shown in Table 3, HCA achieved complete separation (100 % purity) for pure coffee, date pit, and chickpea-adulterated samples. A minor misclassification occurred within the barley-adulterated group, yielding 87.5 % purity. This misclassification likely reflects the intermediate spectral similarity of barley with coffee or other adulterants. Overall, the minimal overlap across clusters highlights HCA's discriminative capability for exploring intrinsic groupings in FTIR data, even without supervised learning. These findings are in agreement with previous reports (e.g., [30]), who also demonstrated the utility of HCA in distinguishing adulterated and non-adulterated food matrices.

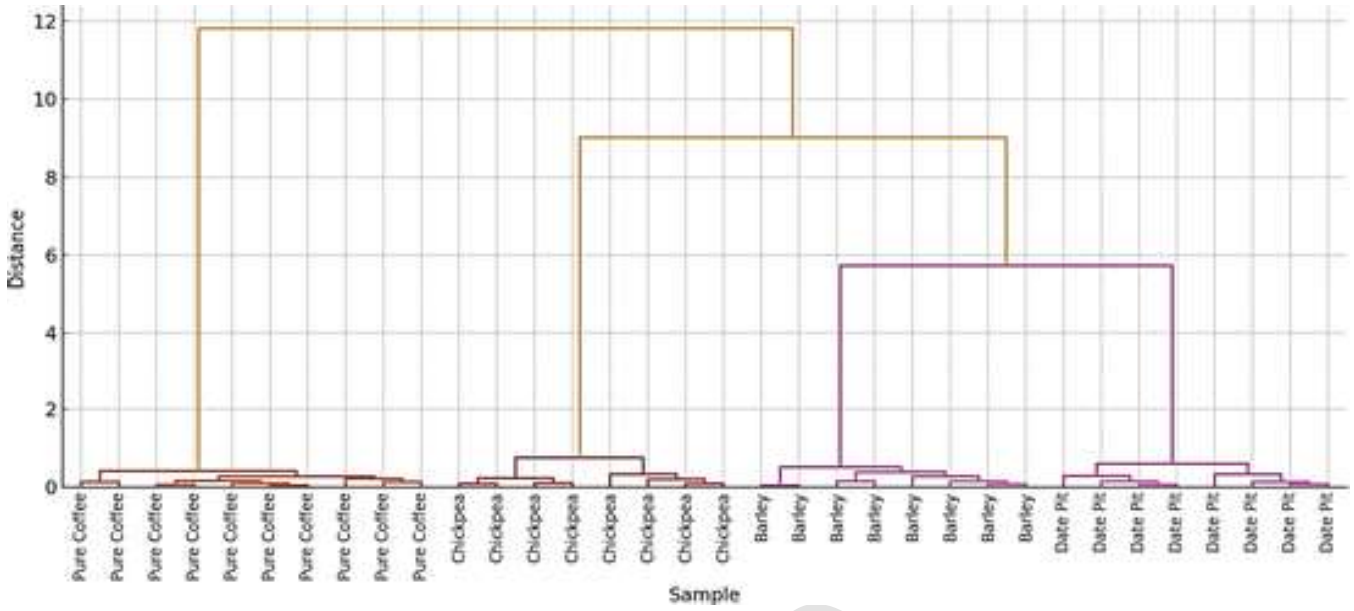

Fig. 5. Dendrogram of hierarchical cluster analysis (HCA) based on FTIR spectral data for classification of pure and adulterated ground coffee samples.

Table 3
Clustering purity results of HCA based on FTIR spectral data.

Cluster ID	Assigned class	No. of samples in cluster	Correctly grouped samples	Cluster purity (%)
Cluster 1	Pure Coffee	10	10	100.0 %
Cluster 2	Date Pit Adulterant	8	8	100.0 %
Cluster 3	Barley Adulterant	8	7	87.5 %
Cluster 4	Chickpea Adulterant	8	8	100.0 %

3.5. Comparative performance evaluation of supervised and unsupervised classification models in coffee adulteration detection

Table 4 presents a comparative analysis of classification accuracy across supervised and unsupervised learning methods for FTIR-based detection of ground coffee adulteration. Among all evaluated models, Hierarchical Cluster Analysis (HCA) outperformed others with an overall accuracy of 96.8 %, achieving 100 % class-wise accuracy for pure coffee, date pit, and chickpea adulteration. While K-Nearest Neighbors (KNN) also showed high performance (91 % overall), especially for pure and date pit samples, it fell short in detecting chickpea adulteration compared to HCA. Random Forest (RF) exhibited good generalization with 83 % overall accuracy but lagged behind HCA in both class-wise accuracy and interpretability. Decision Tree (DT), though simple, delivered the weakest performance with just 70 % accuracy, underscoring its limited capacity to handle the multivariate complexity of FTIR data. The PCA + K-means method, a commonly used unsupervised approach, demonstrated promising results (86.5 % accuracy), yet it failed to match the granularity and precision offered by HCA. These findings emphasize that HCA not only excels in accuracy but also maintains robust class separability and interpretive clarity, making it highly suitable for practical deployment in food fraud detection where labeled datasets are scarce or unavailable.

Table 4
Comparative accuracy of classification models in coffee adulteration detection.

Method	Pure coffee accuracy (%)	Chickpea accuracy (%)	Date pit accuracy (%)	Barley accuracy (%)	Overall accuracy (%)
Random Forest	88	80	85	82	83
KNN	94	86	91	88	91
Decision Tree	78	70	75	72	70
PCA + K-Means	90	85	87	84	86.5
HCA	100	100	100	87.5	96.8

To better contextualize the claimed performance, a comparative summary of representative studies using vibrational spectroscopy for coffee authentication is provided in Table 5. The table contrasts FTIR, NIR, and Raman-based approaches, along with the chemometric methods and performance metrics reported. As seen, our FTIR-based workflow achieves competitive accuracy and, in some cases, superior discrimination power, despite using relatively simple instrumentation and minimal preprocessing. Importantly, our study uniquely demonstrates robust performance across multiple common adulterants (date pit, barley, chickpea), underscoring its practical applicability to routine coffee quality control.

4. Conclusion

This study successfully integrated FTIR spectroscopy with advanced pattern recognition techniques to detect and classify adulteration in ground coffee. The spectral analysis revealed distinct biochemical markers differentiating pure coffee from adulterants, particularly in the lipid ester (1740 cm^{-1}), amide I (1650 cm^{-1}), and carbohydrate (1000–1100 cm^{-1}) regions. Preprocessing methods, including SNV and First Derivative Transformation, improved spectral resolution and model accuracy by up to 9 %. Among supervised models, Random Forest exhibited the strongest performance (83 % testing accuracy), while Decision Tree struggled with complex spectral patterns. Unsupervised learning, particularly HCA, demonstrated exceptional discriminative power (96.8 % accuracy) without requiring labeled training data, making it ideal for preliminary screening. The novelty of this work lies in its comprehensive comparative assessment of supervised and unsupervised approaches within a single framework, clearly demonstrating the

Table 5
Comparison of vibrational spectroscopic approaches for coffee authentication and adulteration detection.

Study	Spectroscopic technique	Data processing/chemometric method	Reported performance	Notes
Wang et al. [32]	FTIR	PCA + LDA	Correct classification >90 %	Authentication of Kona coffee vs. blends
Pauli [33]	FTIR	PCA, LDA	Clear separation of adulterated vs. authentic	Focus on soybean and wheat adulteration
Barrios-Rodriguez et al. [34]	ATR-FTIR	PCA + multivariate models	90–95 % discrimination	Detection of adulterated roasted coffee
de Carvalho Couto et al. [35]	NIR	PLS-DA	>95 % accuracy	Multiple adulterants in Arabica coffee
Sahachairungrueng et al. [36]	NIR hyperspectral + FTIR	PCA + classification	92–96 % accuracy	Differentiation of Arabica and Robusta blends
Fernando et al. [37]	Raman & FTIR (review)	Chemometric models	90–97 % range reported	Overview of vibrational methods for coffee authentication
This study (2025)	FTIR	HCA, PCA + K-means, supervised models	Cluster purity up to 100 %; supervised >95 % accuracy	Robust detection of multiple adulterants (date pit, barley, chickpea)

advantages of HCA and identifying robust FTIR spectral biomarkers for coffee authentication. The findings align with prior research on food fraud detection and confirm that FTIR combined with chemometrics offers a rapid, reliable approach for coffee authentication. Future work should explore real-world validation with multi-adulterant blends and portable FTIR systems for industrial applications. Additionally, future studies should investigate the sensitivity of the proposed approach for trace adulteration (<5 %) to further enhance its applicability in real-world scenarios. It is also important to acknowledge that several challenges remain for large-scale industrial deployment. Variability among coffee products (across brands, origins, and roasting levels), the influence of environmental noise during on-site spectral acquisition, and the complexity of multi-adulterant mixtures may affect robustness. Addressing these factors through larger and more diverse datasets, improved preprocessing strategies, and advanced modeling will be essential to ensure reliable industrial implementation. This study contributes to food safety by providing a framework for combating economically motivated adulteration in the coffee supply chain.

Uncited reference

[38]

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by Ferdowsi University of Mashhad and MULTICAFE Company, Mashhad, Iran (grant number 102538). The funding sources had no involvement that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] V.V. Freitas, L.L.R. Borges, M.C.T.R. Vidigal, M.H. dos Santos, P.C. Stringheta, Coffee: a comprehensive overview of origin, market, and the quality process, Trends Food Sci. Technol. 146 (2024) 104411, https://doi.org/10.1016/j.tifs.2024.104411.
[2] B. Sezer, H. Apaydin, G. Bilge, I.H. Boyaci, Coffee arabica adulteration: detection of wheat, corn and chickpea, Food Chem. 264 (2018) 142–148, https://doi.org/10.1016/j.foodchem.2018.05.037.
[3] J.S. Coqueiro, Beatriz Sales, A. de Lima, J. Cardim de Jesus, R. Rodrigues Silva, S. Passini Barbosa Ferrão, L. Soares Santos, Ensuring authenticity of cinnamon powder: detection of adulteration with coffee husk and corn meal using NIR, MIR spectroscopy and chemometrics, Food Control 166 (2024) 110681, https://doi.org/10.1016/j.foodcont.2024.110681.
[4] D. Suhandy, M. Yulia, S. Widodo, H. Naito, D.F. Al Riza, Fast authentication of Indonesian ground-roasted Arabica coffee adulterated with roasted soybean by portable LED-based fluorescence spectroscopy and chemometrics analysis, Food Chem. 479 (2025) 143791, https://doi.org/10.1016/j.foodchem.2025.143791.
[5] W.L. Cheah, M. Fang, HPLC-based chemometric analysis for coffee adulteration, Foods 9 (7) (2020) 880, https://doi.org/10.3390/foods9070880.
[6] S.S. Nallan Chakravartula, R. Moscetti, G. Bedini, M. Nardella, R. Massantini, Use of convolutional neural network (CNN) combined with FT-NIR spectroscopy to predict food adulteration: A case study on coffee, Food Control 135 (2022) 108816, https://doi.org/10.1016/j.foodcont.2022.108816.
[7] R. Khodabakhshian, H. Seyedalibeyk Lavasani, P. Weller, A methodological approach to preprocessing FTIR spectra of adulterated sesame oil, Food Chem. 419 (2023) 136055, https://doi.org/10.1016/j.foodchem.2023.136055.
[8] M. Masoudi, R. Khodabakhshian, Genetic algorithm-optimized PLS for detecting adulteration in cinnamon powder via FT-IR spectroscopy, Expert Syst. Appl. 290 (2025) 128522, https://doi.org/10.1016/j.eswa.2025.128522.
[9] E.H.D. Paulo, A.M. Rech, F.H. Weiler, M.H.C. Nascimento, P.R. Filgueiras, M.F. Ferrão, Evaluation of adulteration in soy-based beverages by water addition using chemometrics applied to ATR-FTIR spectroscopy, Food Control 166 (2024) 110746, https://doi.org/10.1016/j.foodcont.2024.110746.
[10] S. Shi, K. Zhang, N. Tian, Z. Jin, K. Liu, L. Huang, Y. Jiang, Spectroscopic techniques combined with chemometrics for rapid detection of food adulteration: applications, perspectives, and challenges, Food Res. Int. 211 (2025) 116459, https://doi.org/10.1016/j.foodres.2025.116459.
[11] D. Singh, R. Sharma, S.K. Shivanna, K. Gandhi, R. Singh, T. Sharda, Detection of adulteration in ghee using Raman spectroscopy in combination with chemometrics, Food Control 178 (2025) 111481, https://doi.org/10.1016/j.foodcont.2025.111481.
[12] H. Chen, C. Tan, Z. Lin, Identification of beef adulteration based on near-infrared spectroscopy and an ensemble of radical basis function network, J. Food Compos. Anal. 143 (2025) 107633, https://doi.org/10.1016/j.jfca.2025.107633.
[13] X. Wang, C. Wei, W. Wang, D. Wang, Y. Liu, B. Jia, Y. Jiao, Identification of camellia oil adulteration by using near infrared spectroscopy combined with two-dimensional correlation spectroscopy (2DCOS) analysis, Infrared Phys. Technol. 149 (2025) 105902, https://doi.org/10.1016/j.infrared.2025.105902.
[14] R. Goyal, P. Singha, S.K. Singh, Machine learning-assisted Fourier transform infrared spectroscopy to predict adulteration in coriander powder, Food Chem. 477 (2025) 143502, https://doi.org/10.1016/j.foodchem.2025.143502.
[15] J. Qu, C. Zhang, S. Gao, H. Tian, D. Dong, Integration of near-infrared spectroscopy and comparative principal component analysis for flour adulteration identification, Agric. Commun. 3 (1) (2025) 100073, https://doi.org/10.1016/j.agrcom.2025.100073.
[16] R. Khodabakhshian, M.R. Bayati, B. Emadi, Adulteration detection of Sudan red and metanil yellow in turmeric powder by NIR spectroscopy and chemometrics: the role of preprocessing methods in analysis, Vib. Spectrosc. 120 (2022) 103372, https://doi.org/10.1016/j.vibspec.2022.103372.
[17] J.C.C. Pereira, M.B. Silva, B.D.O. Matos, J.C. de Jesus, R.R.V. Onelli, R.R. Silva, L.S. Santos, Detection of adulteration in cupuaçu pulp using spectroscopy in the infrared in conjunction with multivariate techniques, Food Chem. 478 (2025) 143642, https://doi.org/10.1016/j.foodchem.2025.143642.
[18] L. Hu, S. Ma, C. Yin, Discrimination of geographical origin and detection of adulteration of kudzu root by fluorescence spectroscopy coupled with multi-way pattern recognition, Spectrochim. Acta A Mol. Biomol. Spectrosc. 193 (2018) 87–94, https://doi.org/10.1016/j.saa.2017.12.011.
[19] R. Khodabakhshian, M.R. Bayati, B. Emadi, An evaluation of IR spectroscopy for authentication of adulterated turmeric powder using pattern recognition, Food Chem. 364 (2021) 130406, https://doi.org/10.1016/j.foodchem.2021.130406.
[20] R. Jafari-Marandi, Supervised or unsupervised learning? Investigating the role of pattern recognition assumptions in the success of binary predictive prescriptions, Neurocomputing 434 (2021) 165–193, https://doi.org/10.1016/j.neucom.2020.12.063.

[21] P. Razzaghi, K. Abbasi, J.B. Ghasemi, Chapter 3 – Multivariate pattern recognition by machine learning methods, in: J.B. Ghasemi (Ed.), Machine Learning and Pattern Recognition Methods in Chemistry from Multivariate and Data Driven Modeling, Elsevier, 2023, pp. 47–72.
[22] F.B. de Santana, W. Borges Neto, R.J. Poppi, Random forest as one-class classifier and infrared spectroscopy for food adulteration detection, Food Chem. 293 (2019) 323–332, https://doi.org/10.1016/j.foodchem.2019.04.073.
[23] Q. Liang, Y. Xia, J. Che, Y. Liu, H. Zhang, J. Guo, H. Xue, Detection of water adulteration levels in milk using near-infrared spectroscopy combined with chemometrics, J. Dairy Sci. (2025), https://doi.org/10.3168/jds.2025-26631.
[24] C. Cavdaroglu, N. Altug, A. Serpen, M.H. Öztop, B. Ozen, Comparative performance of artificial neural networks and support vector machines in detecting adulteration of apple juice concentrate using spectroscopy and time domain NMR, Food Res. Int. 201 (2025) 115616, https://doi.org/10.1016/j.foodres.2024.115616.
[25] H.A. Neto, W.L.F. Tavares, D.C.S.Z. Ribeiro, R.C.O. Alves, L.M. Fonseca, S.V.A. Campos, On the utilization of deep and ensemble learning to detect milk adulteration, BioData Mining 12 (2019) 13, https://doi.org/10.1186/s13040-019-0200-5.
[26] A. Sitorus, R. Lapcharoensuk, Exploring deep learning to predict coconut milk adulteration using FT-NIR and micro-NIR spectroscopy, Sensors 24 (7) (2024) 2362, https://doi.org/10.3390/s24072362.
[27] M. Fodor, A. Matkovits, E.L. Benes, Z. Jókai, The role of near-infrared spectroscopy in food quality assurance: a review of the past two decades, Foods 13 (21) (2024), https://doi.org/10.3390/foods13213501.
[28] R. Khodabakhshian, M.H. Abbaspour-Fard, Pattern recognition-based Raman spectroscopy for non-destructive detection of pomegranates during maturity, Spectrochim. Acta A Mol. Biomol. Spectrosc. 231 (2020) 118127, https://doi.org/10.1016/j.saa.2020.118127.
[29] Z. Basati, B. Jamshidi, M. Rasekh, Y. Abbaspour-Gilandeh, Detection of sunn pest-damaged wheat samples using visible/near-infrared spectroscopy based on pattern recognition, Spectrochim. Acta A Mol. Biomol. Spectrosc. 203 (2018) 308–314, https://doi.org/10.1016/j.saa.2018.05.123.
[30] N. Cebi, M.T. Yilmaz, O. Sagdic, A rapid ATR-FTIR spectroscopic method for detection of sibutramine adulteration in tea and coffee based on hierarchical cluster and principal component analyses, Food Chem. 229 (2017) 517–526, https://doi.org/10.1016/j.foodchem.2017.02.072.
[31] M. Hossin, M.N. Sulaiman, A review on evaluation metrics for data classification evaluations, Int. J. Data Mining Knowl. Manage. Process 5 (2015) 01–11, https://doi.org/10.5121/ijdkp.2015.5201.
[32] J. Wang, S. Jun, H.C. Bittenbender, L. Gautz, Q.X. Li, Fourier transform infrared spectroscopy for Kona coffee authentication, J. Food Sci. 74 (5) (2009) C385–C391, https://doi.org/10.1111/j.1750-3841.2009.01173.x.
[33] E.D. Pauli, F. Barbieri, P.S. Garcia, T.B. Madeira, V.R. Acquaro, I.S. Scarminio, S.L. Nixdorf, Detection of ground roasted coffee adulteration with roasted soybean and wheat, Food Res. Int. 61 (2014) 112–119, https://doi.org/10.1016/j.foodres.2014.02.032.
[34] Y.F. Barrios-Rodriguez, Y. Devia-Rodriguez, N.G. Guzmán, Detection of adulterated coffee by fourier-transform infrared (FTIR) spectroscopy associated with sensory analysis, Coffee Sci. 17 (2022), e171970 https://doi.org/10.25186/.v17i.1970, ISSN 1984-3909.
[35] C. Couto, O. Freitas-Silva, E. Oliveira, C. Sousa, S. Casal, Near-infrared spectroscopy applied to the detection of multiple adulterants in roasted and ground Arabica coffee, Foods 11 (2021) 61, https://doi.org/10.3390/foods11010061.
[36] W. Sahachairungrueng, C. Meechan, N. Veerachat, A.K. Thompson, S. Teerachaichayut, Assessing the levels of Robusta and Arabica in roasted ground coffee using NIR hyperspectral imaging and FTIR spectroscopy, Foods 11 (19) (2022) 3122, https://doi.org/10.3390/foods11193122.
[37] D. Fernando, A. Rahmanina, A. Rohman, D. Santosa, Vibrational spectroscopy in combination with chemometrics as an emerging technique in the authentication of coffee, J. Appl. Pharm. Sci. 13 (3) (2023) 012–022, https://doi.org/10.7324/JAPS.2023.130302.
[38] T.A. Teklemariam, F. Chou, P. Kumaravel, J. Van Buskrik, ATR-FTIR spectroscopy and machine/deep learning models for detecting adulteration in coconut water with sugars, sugar alcohols, and artificial sweeteners, Spectrochim. Acta A Mol. Biomol. Spectrosc. 322 (2024) 124771, https://doi.org/10.1016/j.saa.2024.124771.