Agriculture and Food

AIMS Agriculture and Food, 10(3): 577–595.
DOI: 10.3934/agrfood.2025029
Received: 16 March 2025
Revised: 12 June 2025
Accepted: 23 June 2025
Published: 31 July 2025

https://www.aimspress.com/journal/agriculture

Research article

Integration of independently preprocessed spectra to improve classification of repeatedly fried durian chips using portable near-infrared spectrometer and machine learning

Arthit Phuangsombut, Kaewkarn Phuangsombut, Sirinad Noypitak and Anupun Terdwongworakul*

Department of Agricultural Engineering, Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, 73140, Thailand

* **Correspondence:** Email: fengant@ku.ac.th; Tel: +66867751723.

Abstract: Durian chips fried in repeated-use oil can be harmful to health, as prolonged use of the oil leads to excessively high levels of total polar compounds. This study explored the use of a portable near-infrared (NIR) spectrometer combined with machine learning techniques to classify chips into two groups based on frying duration. An alternative fusion technique was proposed by integrating various independently preprocessed spectral data before model analysis. Binary classification was performed to distinguish between chips fried in oil used for 48 hours or less and those fried in oil used for more than that duration. Based on the results, the model developed using partial least squares discriminant analysis with fusion-preprocessed spectral data achieved superior classification performance, with a correlation coefficient of prediction (r_p) of 0.810 and a root mean square error of prediction (RMSEP) of 0.294, compared to the model using traditional preprocessed spectral data (r_p of 0.791 and RMSEP of 0.310). Subsequently, the fusion-preprocessed spectral data were used to develop a support vector machine (SVM) model, which demonstrated strong performance, achieving a precision of 91.9% in minimizing false positives and reducing the misclassification of chips fried in oil for a longer duration compared to those fried for a shorter duration. However, the SVM model lacked robustness when applied to different datasets. When trained and tested on the same dataset, the classification performance was highest for distinguishing between chips fried in oil used for 36 h or less and those fried in oil used for more than 36 h, achieving a precision of 97.8% on the test set. These findings suggested that the combination of a portable NIR spectrometer and the fusion preprocessing technique held promise for rapid, on-site assessment of chips fried in oil used for up to 36 h.

Nonetheless, further studies should be undertaken to enhance the robustness and generalizability of the SVM model.

Keywords: durian chip; fried oil quality; portable device; near-infrared spectroscopy; fusion

1. Introduction

Durian (*Durio zibethinus* L.) is regarded as the most valuable tropical fruit in Southeast Asia, where Thailand leads in the fruit's export value. At the time of harvest, durian fruit can be over-supplied. Therefore, farmers commonly process the excess fruit into fried durian chips that can be exported to countries that do not allow fresh imports of durian, thus expanding the export market for durian.

While frying at high temperature, oxygen in the air and moisture in the food lead to thermal polymerization, oxidation, and hydrolysis as a result of complex reactions [1]. Consequently, the chemical and physical properties of the frying oil will change greatly. Furthermore, the oil quality will deteriorate due to the frying process generating harmful decomposition products [2]. Total polar compounds (TPCs) are formed during frying and increase with frying time; thus, this parameter is normally used to assess the quality of the frying oil [3,4,5]. A higher TPC content from a longer frying time is harmful to health, and so it has been limited to a maximum of 25% in some countries [6]. However, laboratory determination of the TPC content is time consuming.

Near-infrared spectroscopy (NIRS) has the advantages of being a nondestructive, rapid, and environment-friendly method that has been applied to evaluate the quality of frying oil. Some studies reported strong correlations between total polar materials and NIRS in soy-based frying oils [7,8] and in refined hazelnut and peanut frying oils [9]. Visible spectroscopy (Vis) and NIRS were studied by Kazemi et al. [10] to evaluate changes in the TPC content from frying canola oils. As the TPC content is linearly related to the frying time, most studies have predicted the frying time in reference to the TPC content. Liu et al. [11] achieved good results in applying NIRS to predict frying times of soybean oil. Successful prediction of the frying time of soybean oil has extended to other oils. For example, Liu et al. [12] successfully used NIRS to evaluate the frying-oil quality of two edible oils (rapeseed oil and peanut oil) based on their frying times; the frying time was accurately predicted based on the NIR spectral output in the range 400–2500 nm and the first derivative pre-processing. Liu et al. [13] established a model using Bayesian ridge regression based on selected wavelengths of the NIR spectra of rapeseed oil, peanut oil, and soybean oil, with prediction performance values of 0.985 and 0.546 being reported as the coefficient of determination (R^2) and the root mean square error of prediction, respectively. Another study used NIRS to rapidly determine the degradation from frying rapeseed oil at three different temperatures, with high accuracy in TPC prediction being reported with R^2 and standard error of cross-validation values of 0.98 and 1.25%, respectively [14]. Additionally, several studies have focused on predicting key quality parameters of oils and biodiesel using NIRS. These include the cetane number in waste cooking oil-derived biodiesel [15], the peroxide value and free fatty acid content of cooking oil [16], and the anisidine value of oil during crispy pork rind frying [17].

NIRS has been applied to evaluate the quality of different types of chips. Palarea-Albaladejo et al. [18] estimated the relative content of fat in potato chip products using NIRS as a rapid and non-destructive technique. Pedreschi et al. [19] investigated the feasibility of implementing on-line NIRS to monitor the oil content in fried potato chips, with a prediction model developed based on NIR

interactance spectra and partial least squares regression (PLSR) being highly accurate. Areekij et al. [20] determined the fat content in deep-fried taro chips using NIRS in combination with PLSR. NIR hyperspectral imaging, which provided additional spatial data over spectral data, combined with a partial least squares discriminant analysis was used to classify potatoes into cooked or fried [21].

Emerging data fusion, which integrates multi-sourced data, provides more potentially detailed complementary information than relying on a single source data [22]. This technique optimizes and integrates different sources of data to acquire more reliable, comprehensive, and rich feature information. The integrated data can be used to build regression or classification models for analysis of samples, with improved model performance. There are three main types of data fusion: low-level, mid-level, and high-level [23]. In low-level (data-level) fusion, raw data from different sources are directly combined into a single matrix. Mid-level (feature-level) fusion involves extracting relevant features from each dataset and merging them into a unified feature matrix. High-level (decision-level) fusion combines the individual classification results from different models or datasets to produce a final joint decision.

Xia et al. [24] demonstrated that integrating NIRS and electronic nose (E-nose) technologies using a feature-level fusion strategy offered a rapid, accurate, and non-destructive method for assessing overall tea quality. Important features from the data obtained using NIRS and E-nose were extracted and fused. The best performance for 6 classifications of 98.13% was achieved based on a support vector machine approach. Hong et al. [25] used mid-level data fusion for datasets using rapid evaporative ionization mass spectrometry and inductively coupled plasma mass spectrometry to develop classification model for salmon origin and production methods, including using principal component analysis (PCA) to reduce data dimensionality before fusion. They reported that the mid-level data fusion coupled with partial least squares discriminant analysis was best at improving the classification. Xu et al. [26] investigated maize seed disease identification using multi-source spectral information fusion techniques based on decision-level fusion, where the classification results from the different datasets or models were combined to generate a final, joint inference outcome. They reported that the decision-level fusion of spectral and image data from hyperspectral imaging (HSI) provided the best accuracy (98.12%). An et al. [27] also used decision-level fusion to combine NIR, mid-infrared, and Raman spectral data for accurate quantification of the staling degree of Chinese steamed bread.

Data fusion has been applied to integrate Vis/NIR hyperspectral imaging and Fourier transform-NIR spectroscopy data of kiwifruit for improved prediction of soluble solids content, flesh firmness, and dry matter [28]. Ping et al. [29] reported the development of a non-destructive model to identify the origin of ginseng samples based on data from hyperspectral imaging in conjunction with X-ray technology. Their model used the fusion of hyperspectral spectral data and X-ray image texture data to produce substantially better performance than the model using only hyperspectral spectral information. In addition, the fusion of Vis-NIR and NIR hyperspectral imaging information has been used to qualitatively evaluate multiple qualities in chicken [30] and to predict norfloxacin residues in mutton [31]. Posom et al. [32] proposed an alternative data fusion approach that integrated preprocessing techniques within an NIR spectrum, with multi-blocks of the spectral preprocessing technique applied to the NIR spectra that resulted in a better prediction model for online measurement of the moisture content and higher heating value of sugarcane bagasse.

All these studies focused on the evaluation of frying oil quality. However, chips as the end product, have not been investigated based on frying using low quality oil with a long frying time. Thus, the current research aimed to apply NIRS for the classification of the quality of durian chips fried in oils

with different frying times. Furthermore, most other investigations have focused on integration of information from different instruments to achieve data fusion, with only a few studies investigating the integration of preprocessing across the entire NIR spectrum. Therefore, the current study also explored an alternative technique for fusing independently preprocessed NIR spectra based on measurements from a miniature spectrometer. The classification accuracy was compared between models built using fused independently preprocessed NIR spectra and those constructed with individually preprocessed NIR spectra.

2. Materials and methods

2.1. Durian chip samples

Durians of uniform ripeness were harvested from a single orchard in Chanthaburi province, Thailand. Their stone-free pulp was extracted and sliced into approximately 3 mm thick chips for deep-frying. Then, these chips were sealed in plastic bags and refrigerated at 8 ± 2 ℃. Before each frying experiment, 150 g batches of durian chips were removed from refrigeration and allowed to equilibrate to ambient temperature for 1 h. Palm oil purchased from a local supermarket was used as the frying oil. Frying was carried out in a stainless-steel electric fryer (model TF-40A; Taiji & Co., Ltd.; Kanagawa, Japan). The frying process occurred at 180 ± 5 ℃ using 2 L of oil. Once the desired temperature of oil had been reached (which took about 5 min), each batch of durian chips was fried for 10 min. Then, the chips were removed in a strainer from the oil and the heating was halted to allow the oil to cool for at least 10 min to complete one round of the experiment. Eight rounds of experiments were evenly spaced during 8 h each day and this schedule continued for 10 consecutive days. Nine durian chip samples from each round or each hour were randomly collected for acquisition of NIR measurement. In total, 720 samples of durian chips were obtained for further experimentation.

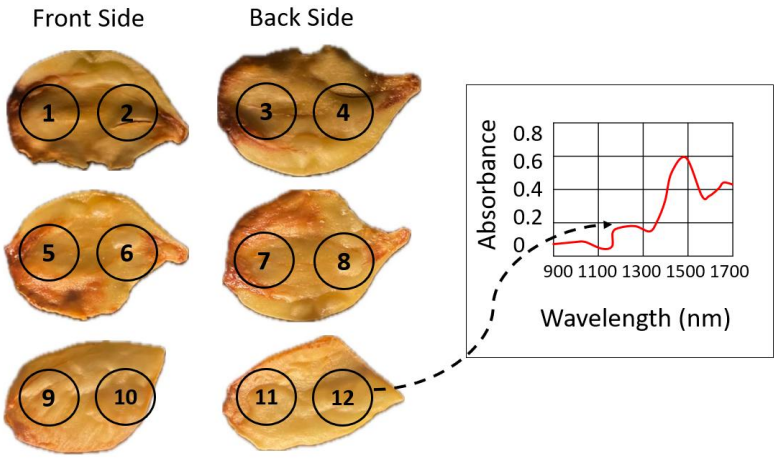


Figure 1. Scan positions of front side and back side on three durian chip samples.

2.2. Spectral acquisition

Each durian chip was subjected to NIR acquisition using a spectrometer (DLP NIR scan Nano;

Texas Instruments; Dallas, TX, USA) with a detector made of an In-Ga-As linear array. Prior to starting a scan and following every 10 scans of samples, the white reference was acquired by scanning a spectralon bar (100 × 25 × 10 mm in length, width, and thickness, respectively; Simtrum; Singapore). Each chip was scanned in diffuse reflectance mode in the wavelength range 900–1700 nm with an integration time of 4.6 s and a 3.2 nm increment. Configuration, absorbance recording, and absorbance file saving were managed using the NIRscan Nano GUI software (Texas Instruments; Dallas, TX, USA). Each side of the chip was scanned at two positions and the four spectra were averaged to represent the spectrum of the chip (Figure 1).

2.3. Development of classification model

2.3.1. Partial least squares discriminant analysis

Initially, partial least squares discriminant analysis (PLS-DA) was applied to develop classification models based on the NIR spectra of the intact chip samples using the MATLAB software (version 2021a; MathWorks; Natick, MA, USA). Six samples from the nine chip samples of each round were randomly allocated to a training set and the remaining three samples to a test set. In total, the training and test sets consisted of 480 and 240 samples, respectively. The models were built to classify the samples into 10 groups of frying time, five groups of frying time, and two groups of frying time. For the 10 and five groups, there was equal separation of the number of frying hours. For the two groups, the number of frying times at 48 h was used to separate the samples into each group. The 48-hour limit was based on findings indicating that oil heated for more than 48 h would have total polar compounds exceeding 25%, rendering it unsuitable for frying according to the regulations in most countries [10]. For model training using the training dataset, leave-one-out cross validation was used to determine the optimal number of factors for the PLS-DA classification model. Subsequently, the optimal PLS-DA model was applied to classify the samples in the test set. The performance of each model was evaluated based on the correlation coefficient of prediction (r_p), the root mean square error of calibration (RMSEC), and the root mean square error of prediction (RMSEP).

The NIR spectra were prepared for model analyses using various preprocessing methods: original spectra; spectra preprocessed using the standard normal variate (SNV); spectra preprocessed using the second derivative (2D) with eleven data points; spectra preprocessed using the first derivative (1D) with nine data points; spectra preprocessed using SNV and 2D (SNV-2D); and spectra preprocessed with SNV and 1D (SNV-1D). Prior to the analyses, all forms of spectra were subjected to a genetic algorithm (GA) to search for feature wavelengths [33]. GAs are bio-inspired optimization methods used for feature selection that mimic the process of natural evolution by applying operations, such as selection, crossover, and mutation, to iteratively refine a population of candidate solutions. In the context of feature selection, each candidate solution (chromosome) represents a subset of features. The algorithm's objective is to identify the minimal subset of features that maximizes the performance of a machine learning model [34].

Two fusion strategies were utilized. In the first strategy, the original spectra were combined directly with all preprocessed spectra, with the combined spectra then being subjected to GA analysis (GA-COM) to select feature wavelengths for further study. In the second strategy, the original spectra and all preprocessed spectra were first analyzed individually using GA for feature selection, followed by integration for further analysis (COM-GA).

2.3.2. Classification analysis using support vector machine

Machine learning techniques were also applied to construct the classification models. The classification model was developed using the classification learner toolbox of the MATLAB software (R2021a; Mathworks; Natick, MA, USA). The support vector machine (SVM) method has been reported as an optimal machine learning method for durian [35,36]. Therefore, SVM was selected as the exclusive classifier for comparison purposes. SVM classification is a supervised machine learning algorithm designed to fit the best possible line within a predefined margin of error (epsilon). It achieves this by identifying the hyperplane that best represents the data points in a continuous space, while maximizing the margin (distance) between the closest data points, known as support vectors, and the hyperplane. The primary objective of SVM classification is to minimize the error for data points that fall outside the margin [37]. In the current study, the spectral data in the training set was used to train the models. To prevent overfitting, the K-fold cross-validation approach was applied during model development. In the classification learner toolbox, 5-fold cross-validation was used as the default setting. The classification models were evaluated and compared using the confusion matrix, based on accuracy, recall rates, precision, and F1-score, as defined in Eqs (1)–(4), as well as the receiver operating characteristic (ROC) curve and the area under the ROC curve (ROC-AUC).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \tag{1}$$

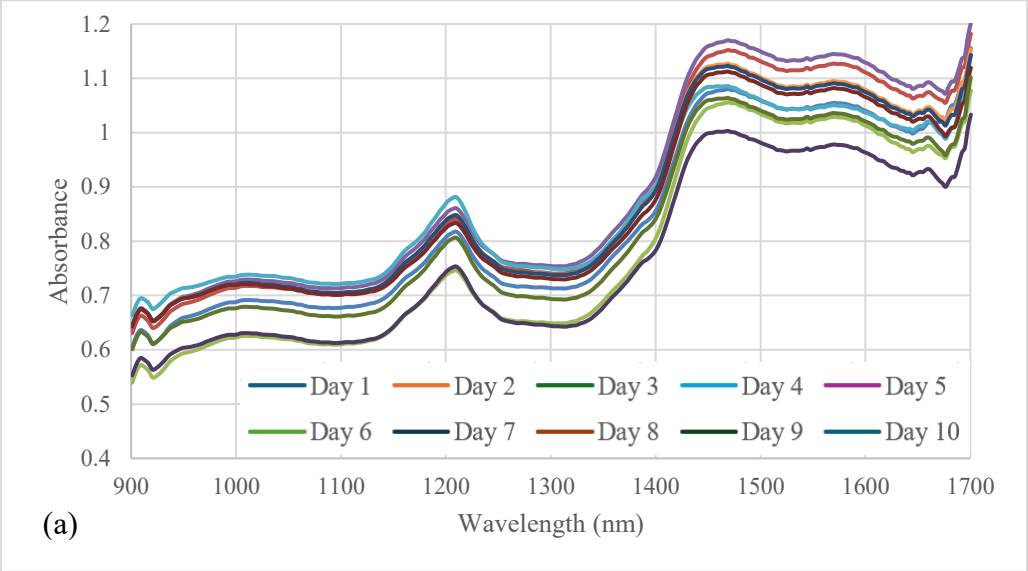
$$Recall = \frac{TP}{TP+FN}, \tag{2}$$

$$Precision = \frac{TP}{TP+FP}, \tag{3}$$

$$F1\text{- }Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{4}$$

where TP, TN, FP, and FN are the true positive class, true negative class, false positive class, and false negative class, respectively.

Accuracy measures how often a classification model correctly predicts the overall outcomes. It is expressed either as a value in the scale range 0–1 or as a percentage. Precision evaluates a model's ability to correctly identify positive cases when predicting the positive class, reflecting its capacity to avoid labeling negative samples as positive. Similar to accuracy, precision is also measured using the scale range 0–1 or as a percentage. Recall assesses a model's ability to correctly identify positive instances from all actual positive samples in the dataset. Like accuracy and precision, recall is measured using the scale range 0–1 or as a percentage. The F1-score balances recall (correctly predicted actual positives) and precision (actually positive predicted positives), which is useful for imbalanced data. The ROC graphically shows a binary classifier's diagnostic ability across different decision thresholds by plotting the true positive rate against the false positive rate. The ROC-AUC is a single value in the range 0–1 summarizing overall performance, with higher values indicating better classification [38].



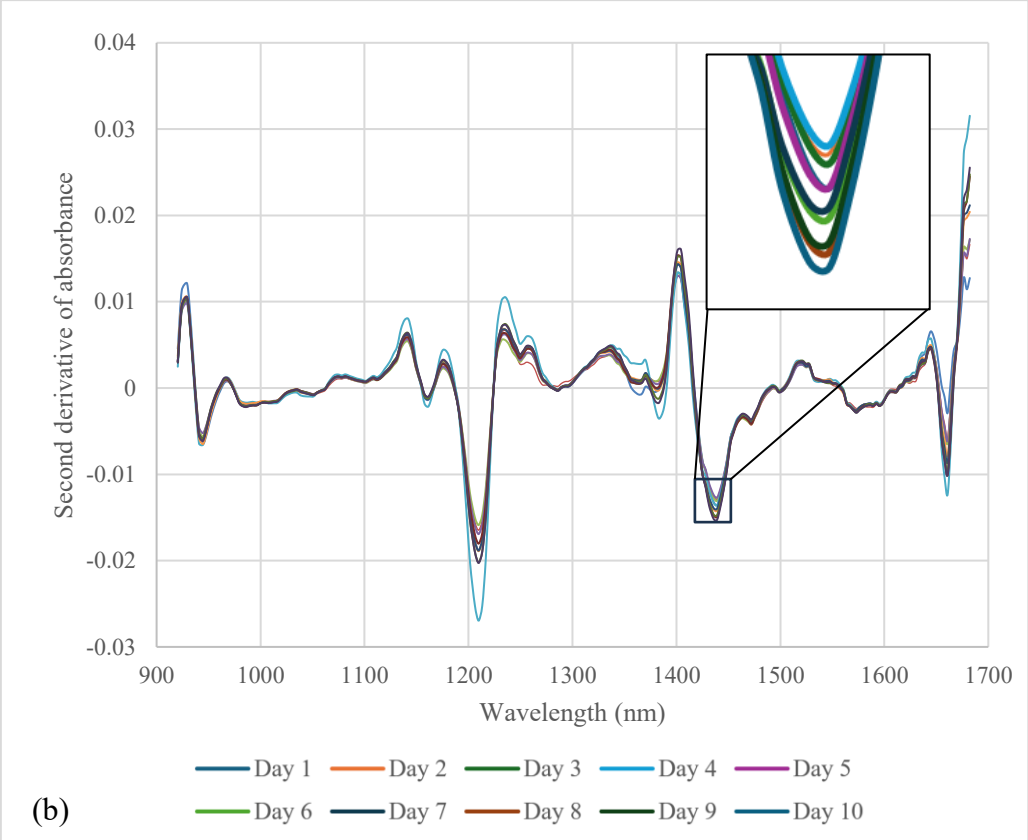


Figure 2. (a) Absorbance of durian chip samples for different frying times (averaged over 8 h for each day) and (b) Second-derivative preprocessed absorbance of durian chip samples for different oil frying times, with inset highlighting differences around 1440 nm.

3. Results and discussion

3.1. Spectral characteristics and sample variability

The average NIR spectra at various frying times are shown in Figure 2(a). A distinct peak, centered around 1200 nm, represented the primary NIR absorption band in oils or fats, corresponding to the CH2 second overtone [32]. A prominent water absorption band was clearly visible in the 1440–1450 nm region, associated with the first overtone of O-H stretching vibration [39]. In the original spectra, no clear relationship with oil frying time was observed, as absorbance changes did not consistently correlate with increased frying duration. To improve peak resolution, second derivative (2D) preprocessing was applied to the spectra. A notable change in 2D absorbance—partially influenced by oil frying time—was observed around 1440 nm (Figure 2[b]). The frying durations could be grouped into two categories: 1–5 days and 6–10 days. In durian chips, the peak at 1440 nm may have reflected variations in moisture content or water activity, which were affected by the dehydration process during frying and could vary with the duration of oil use.

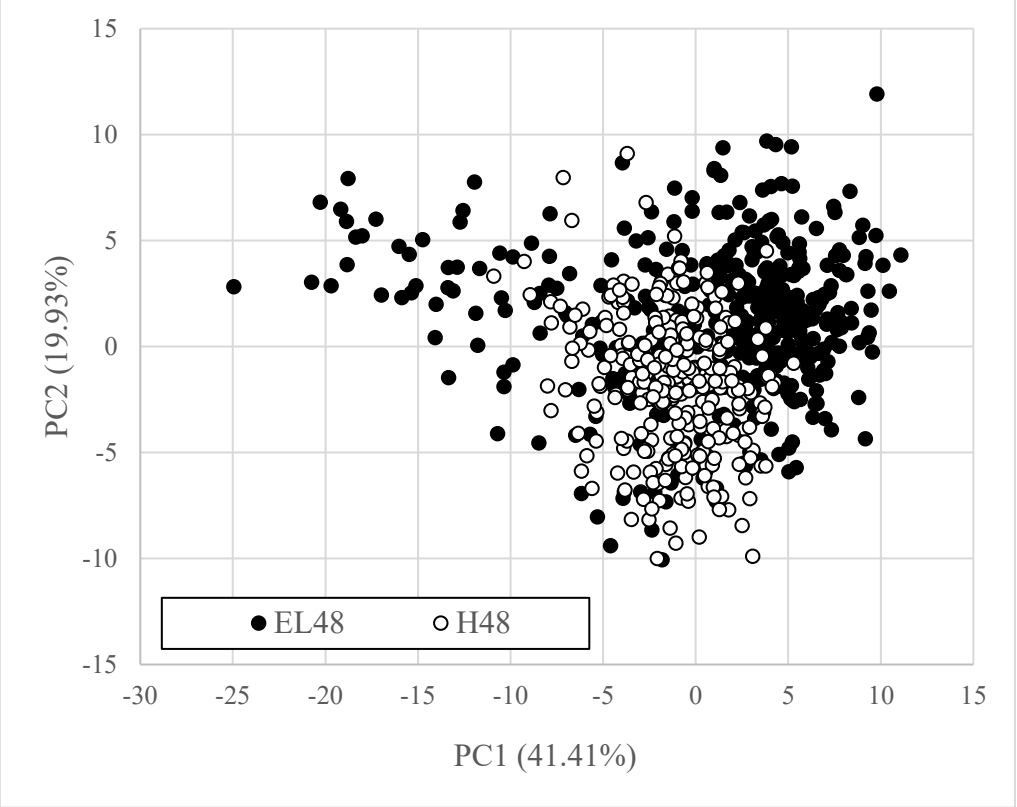


Figure 3. Score plot of first two principal components (PC1 and PC2), illustrating the distribution of chip samples. Black circles represent chips fried in oil used for 48 h or less (EL48), while white circles indicate chips fried in oil used for more than 48 h (H48).

PCA was performed on the original NIR absorbance spectra to visualize the structural relationships between sample groups: chips fried in oil used for 48 h or less (EL48-chips) and those fried in oil used for longer than 48 h (H48-chips). The resulting score plot (Figure 3) revealed that the

first principal component (PC1) explained 41.41% of the variance and the second principal component (PC2) accounted for 19.3%. Despite all samples clustering together, indicating a common origin, the EL48-chips had a wider spread, suggesting greater variability in their NIR absorbance. Conversely, the H48-chips formed a tighter cluster, implying more similar NIR characteristics in chips fried in older oil. Given that these two PCs captured only 61.34% of the variance and PCA is an unsupervised technique, a clear separation between the groups was not achieved. However, the potential for improved classification existed by applying preprocessing techniques to the NIR spectra and using a supervised algorithm.

3.2. Classification performance

3.2.1. Partial least squares discriminant analysis

Each sample in the training and test sets used to construct the classification models was assigned a class value of 1 if the chips were fried in oil used for 48 h or less and a value of 2 if fried in oil used for more than 48 h. PLS-DA classification models based on various forms of NIR absorbance were developed; their performance was compared for two-group classification, as shown in Table 1. Before using the spectra for model development, they were preprocessed and analyzed to identify feature wavelengths using GA. When classifying chip samples into two groups based on the duration of oil usage for frying, the model using GA-1D spectra achieved the best performance, with r_p and RMSEP values of 0.790 and 0.310, respectively. Notably, the model based on spectra preprocessed using two methods did not perform better than the model based on spectra processed using a single method.

Two fusion strategies were explored to improve classification performance. The first strategy, referred to as GA-COM, involved combining all preprocessed spectra. Then, the resulting combined spectra were analyzed using GA to extract feature wavelengths prior to model development. In the second strategy (COM-GA), GA-selected features from the original spectra and all preprocessed spectra were combined and subsequently used to build the model. As presented in Table 2, the GA-COM model outperformed the COM-GA model, achieving r_p and RMSEP values of 0.810 and 0.294, respectively. Furthermore, the GA-COM model had superior performance compared to the GA-1D model, which was the best-performing model among those based on non-fusion spectra. This result partially aligned with the findings of Posom et al. [32], who reported that models utilizing wavelength multi-blocks of the spectral preprocessing technique on full spectra provided better predictions of the moisture content and higher heating value of cane bagasse than models based on raw spectra or preprocessed full spectra. However, the multi-block technique still requires the optimization of block size within the spectrum for each preprocessing method to achieve the optimal model.

Figure 4 presents the regression coefficients of the GA-COM model as a function of feature wavelength. The highest regression coefficient was at 1311 nm, followed by 933 nm and 1442 nm, respectively. The NIR absorption band near 1311 nm is typically associated with the first overtone of C–H stretching vibrations and is commonly observed in lipid-related compounds [40]. The absorption around 933 nm is often attributed to combination bands involving C–H stretching and bending vibrations, particularly in vinyl groups [40]. The band at 1442 nm corresponds to one of the prominent peaks observed in the D2 spectra shown in Figure 2(b).

Table 1. Statistical performance parameters of PLS-DA classification models for durian chips fried in oil used for different durations, grouped into two categories. The models were developed using preprocessed spectra with feature selection performed via a genetic algorithm approach.

Preprocessing technique	No. of factors	r_c	RMSEC	r_p	RMSEP
No preprocessing (O)	15	0.845	0.267	0.777	0.321
Standard normal variate (SNV)	17	0.851	0.262	0.773	0.318
Second derivative (2D)	15	0.826	0.282	0.749	0.333
First derivative (1D)	14	0.829	0.280	0.790	0.310
Standard normal variate and second derivative (SNV-2D)	12	0.794	0.304	0.738	0.341
Standard normal variate and first derivative (SNV-1D)	14	0.823	0.284	0.777	0.318

Note: r_c = Correlation coefficient of calibration, r_p = Correlation coefficient of prediction, RMSEC = Root mean square error of calibration, and RMSEP = Root mean square error of prediction.

To further analyze the data and compare classification performance with the GA-COM approach, a multi-block preprocessing technique was applied—originally proposed by Posom et al. [32] for spectral data from single-instrument data fusion. The spectral data were segmented into four blocks of equal wavelength range. Then, by iteratively combining the original, 1D, 2D, and SNV preprocessed spectra into these blocks, 256 distinct multi-preprocessed spectra were generated. Each of these was used to construct a PLS-DA model, resulting in 256 models. The optimal model from this set of 256 had predictive performance (r_p = 0.808, RMSEP = 0.294, Table 2) comparable to GA-COM, with its four blocks consisting of the original, 1D, 1D, and 2D spectra. Notably, this multi-block technique achieved similar performance with fewer latent variables (20) than GA-COM (27). However, the proposed GA-COM technique offered an alternative approach for data fusion within spectra from a single instrument.

Table 2. Statistical parameters of classification performance of partial least squares discriminant analysis models for durian chips fried in oil used for different durations based on two groups using different fusion techniques of preprocessed spectra.

Preprocessing fusion technique	No. of factors	r_c	RMSEC	r_p	RMSEP
Combination of GA-selected features from O, SNV, 2D, 1D, SNV-2D, and SNV-1D (COM-GA)	19	0.877	0.240	0.786	0.347
GA-selected features from a combination of O, SNV, 2D, 1D, SNV-2D, and SNV-1D spectra (GA-COM)	27	0.850	0.263	0.810	0.294
Multi-block technique	20	0.856	0.259	0.808	0.294

Note: No preprocessing (O); Standard normal variate (SNV); Standard normal variate and second derivative (SNV-2D); Standard normal variate and first derivative (SNV-1D); Combination (COM); Genetic algorithm (GA); r_c = Correlation coefficient of calibration; r_p = Correlation coefficient of prediction; RMSEC = Root mean square error of calibration; and RMSEP = Root mean square error of prediction.

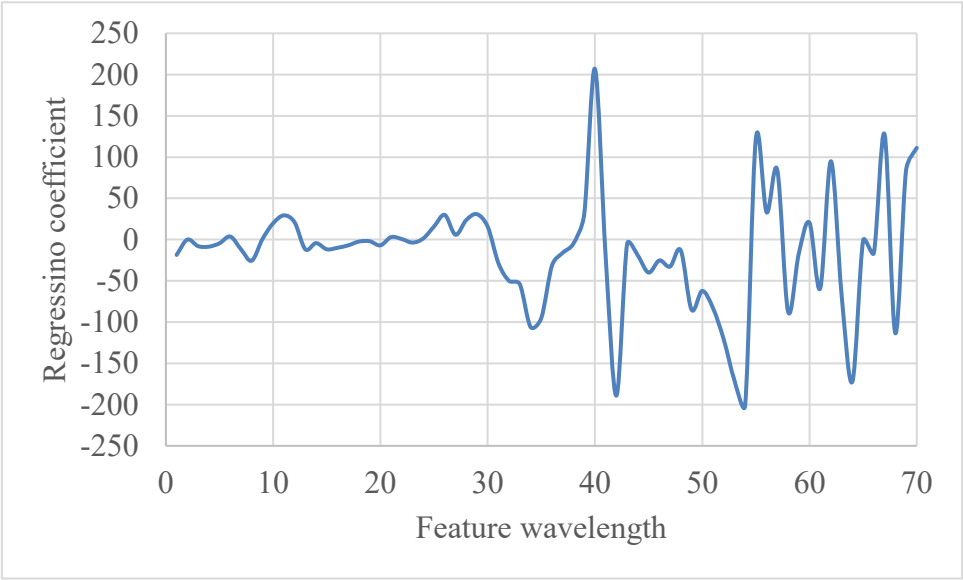


Figure 4. Regression coefficients of PLS-DA model based on GA-COM spectra.

3.2.2. Support vector machine classification

SVM classification, a machine learning technique, was applied to the GA-COM processed spectra for binary classification and to evaluate the model's generalization ability. The data were structured into three binary datasets based on frying oil usage time: 48 h, 36 h, and 24 h. For the 48-hour criterion, the groups were defined as durian chips fried in oil used for 48 h or less (EL48, the positive class) and those fried in oil used for longer than 48 h (H48, the negative class). Similarly, the 36-hour criterion yielded groups EL36 (≤ 36 h) and H36 (> 36 h), and the 24 h criterion resulted in groups EL24 (≤ 24 h) and H24 (> 24 h).

Two analyses were conducted. First, an SVM classification model was developed using EL48 and H48 training data. Then, this model was used to predict data from the following test sets: EL48 and H48, EL36 and H36, and EL24 and H24. In the second analysis, the model was developed using a training set and tested with the corresponding test set. The first model was trained using EL48 and H48 data from the training set and evaluated using the corresponding EL48 and H48 data from the test set. Notably, the first model was the same as that in the first analysis. Similarly, separate models were created and evaluated using EL36 and H36 data for the second model and EL24 and H24 data for the third model.

In the analyses, the positive class was defined as durian chips from the EL48, EL36, and EL24 datasets, while the negative class corresponded to durian chips from the H48, H36, and H24 datasets, respectively. The confusion matrices from the first and second analyses are shown in Figures 5 and 6, respectively, while the classification results are presented in Tables 3 and 4. Additionally, the ROC curves for the first and second analyses are displayed in Figures 7 and 8, respectively.

In the first analysis, the first model was trained and tested on the EL48+H48 dataset and performed relatively well on this dataset. For the test data, it achieved an accuracy of 87.5%, which was considered high (Figure 5[a] and Table 3). Given the class imbalance, the precision and recall were more appropriate performance metrics. In this study, minimizing false positives was a key priority—specifically, instances where EL48 was predicted while the actual class was H48. Therefore, precision

was the key metric, with a high precision value indicating a low number of false positives. The model performed well in this regard, as reflected by the precision of 91.9% (Table 3). A slightly low recall of 86.8% was acceptable in this case, since minimizing false positives was the priority. The F1-score was 89.3%, indicating a strong balance between precision and recall. Furthermore, the ROC curve (Figure 7[a]) and the high AUC value of 0.953 supported the strong performance of the SVM model in this self-testing scenario for class separation (trained and tested on EL48+H48).

To evaluate the model's generalization capability using different data sets, the EL48+H48 model was tested on the EL36+H36 and EL24+H24 datasets. When applied to the EL36+H36 dataset, the model exhibited a performance decline across nearly all metrics. Accuracy dropped to 83.3%, precision fell considerably to 75.0%, recall slightly increased to 94.4%, F1-score declined to 83.6%, and AUC decreased to 0.922 (Figure 5[b], Figure 7[b] and Table 3). This performance indicated that the patterns learned from the EL48+H48 data were not fully transferable to the EL36+H36 data. This discrepancy could have been due to some differences in the NIR characteristics of the durian chips fried in oil used for different durations that the model struggled to handle. The high recall but lower precision suggested the model effectively identified most actual positive cases in the EL36+H36 dataset but also produced a higher number of false positives.

		Predicted	
		EL48	H48
Actual	EL48	125	19
	H48	11	85

(a)

		Predicted	
		EL36	H36
Actual	EL36	102	6
	H36	34	98

(b)

		Predicted	
		EL24	H24
Actual	EL24	69	3
	H24	67	101

(c)

Figure 5. Confusion matrices of test data using SVM model trained on EL48+H48 data and tested on (a) EL48+H48 data, (b) EL36+H36 data, and (c) EL24+H24 data.

Table 3. Classification results of SVM models trained on EL48+H48 and tested on datasets: EL48+H48, EL36+H36, and EL24+H24.

	EL48+H48 Training	EL48+H48 Test	EL36+H36 Test	EL24+H24 Test
Accuracy (%)	88.8	87.5	83.3	70.8
Precision (%)	92.1	91.9	75.0	50.7
Recall (%)	88.9	86.8	94.4	95.8
F1-Score (%)	90.5	89.3	83.6	66.3
ROC (AUC)	0.950	0.953	0.922	0.862

Note: EL48 = Durian chips fried in oil used for 48 h or less; H48 = Durian chips fried for longer than 48 h; EL36 = Durian chips fried in oil used for 36 h or less; H36 = Durian chips fried for longer than 48 h; EL24 = Durian chips fried in oil used for 24 h or less; H24 = Durian chips fried for longer than 24 h.

Based on generalization using the EL24+H24 dataset, the model's performance deteriorated even further. Accuracy dropped to 70.8%, precision fell to 50.7%, recall remained very high at 95.8%, the F1-score decreased substantially to 66.3%, and the ROC (AUC) declined to 0.862 (Figure 5[c], Figure 7[c] and Table 3). These results indicated that the differences between the EL48+H48 and EL24+H24 datasets were more pronounced, which made it more difficult for the model to generalize effectively. The very low precision reflected a high rate of false positive predictions. Although the model successfully identified most of the actual positive cases (as reflected in the high recall), it frequently misclassified negative cases as positive.

Thus, the SVM model demonstrated a lack of robustness when applied to different datasets. The high recall coupled with low precision in the generalization tests suggested a tendency toward overpredicting the positive class when encountering unfamiliar data. Additionally, the AUC scores consistently declined as the test data became more dissimilar from the training data, indicating a reduced ability to effectively distinguish between the two classes. In future investigation, exploring different machine learning models might improve generalization and it would be interesting to get a more comprehensive understanding of the data and model behavior with models trained on EL3+H36 and EL24+H24 and tested on other datasets.

			Predicted	
			EL48	H48
Actual		EL48	125	19
		H48	11	85

(a)

		Predicted	
		EL36	H36
Actual	EL36	91	17
	H36	2	130

(b)

		Predicted	
		EL24	H24
Actual	EL24	52	20
	H24	12	156

(c)

Figure 6. Confusion matrices of test data for SVM models: (a) trained and tested on EL48+H48 data, (b) trained and tested on EL36+H36 data, and (c) trained and tested on EL24+H24 data.

In the second analysis, the SVM models were trained and tested on the same datasets. The model based on the EL48+H48 data performed well in minimizing false positives (Figure 6[a], Table 4 and Figure 8[a]), demonstrating identical performance to that observed in the first analysis. The El36+H36 model had the highest performance among the three. It achieved over 92% accuracy on both training and test sets (Figure 6[b] and Table 4). The precision was very high (97.8%), indicating a low rate of false positives, ideal for applications where avoiding false positives was critical. The model misclassified very few H36 samples as EL36. The slight trade-off was in recall (84.3%), which meant it missed some actual EL36 samples. The F1-score of 90.5% meant that overall performance was strong,

with a good trade-off between catching true positives and avoiding false positives. The ROC (AUC) of 0.971 was the highest, indicating the most effective discrimination between EL36 and H36 data (Figure 7(b) and Table 4).

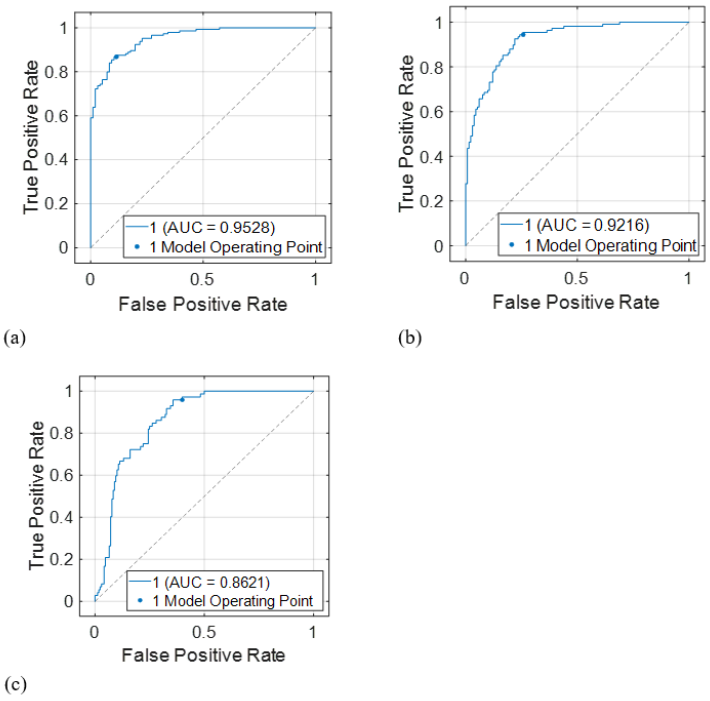


Figure 7. RUC curve of test data using SVM model trained on EL48+H48 data and tested on (a) EL48+H48 data, (b) EL36+H36 data, and (c) EL24+H24 data.

The model trained and tested on the EL24+H24 dataset had the lowest overall performance among the three configurations. Although the training accuracy was relatively high at 90.4%, the test accuracy declined to 86.7% (Figure 6[c] and Table 4). Test precision was moderate at 81.2%, while recall dropped notably from 85.4% in training to 72.2% in testing. Consequently, the model achieved the lowest test F1-score (76.5%), suggesting a moderate balance between precision and recall, but without strong performance in either. Despite a strong ROC AUC of 0.926 (Figure 7[c] and Table 4), the relatively modest precision indicated that the model still produced a considerable number of false positives.

Thus, in the second analyses, the EL36+H36 model achieved the best performance, particularly in precision, among the three configurations, making it the preferred model when the cost or consequence of a false positive was higher than that of a false negative. The best performance was possibly due to more distinct and less overlapping spectral characteristics, more balanced class distribution, and lower within-class variance. Furthermore, with a slightly lower training precision (96.9%) compared to the test precision (97.8%), the model may have avoided overfitting and instead generalized well to unseen data—reflected in its high AUC (0.971) and F1-score (90.5%).

Table 4. Classification results of SVM models trained and tested on same datasets: EL48+H48, EL36+H36, and EL24+H24.

	EL48+H48		EL36+H36		EL24+H24	
	Training	Test	Training	Test	Training	Test
Accuracy (%)	88.8	87.5	92.9	92.1	90.4	86.7
Precision (%)	92.1	91.9	96.9	97.8	83.1	81.2
Recall (%)	88.9	86.8	87.0	84.3	85.4	72.2
F1-Score (%)	90.5	89.3	91.7	90.5	84.2	76.5
ROC (AUC)	0.950	0.953	0.968	0.971	0.958	0.926

Note: EL48 = Durian chips fried in oil used for 48 h or less, H48 = Durian chips fried in oil used for longer than 48 h; EL36 = Durian chips fried in oil used for 36 h or less; H36 = Durian chips fried for longer than 48 h; EL24 = Durian chips fried in oil used for 24 h or less; H24 = Durian chips fried for longer than 24 h.

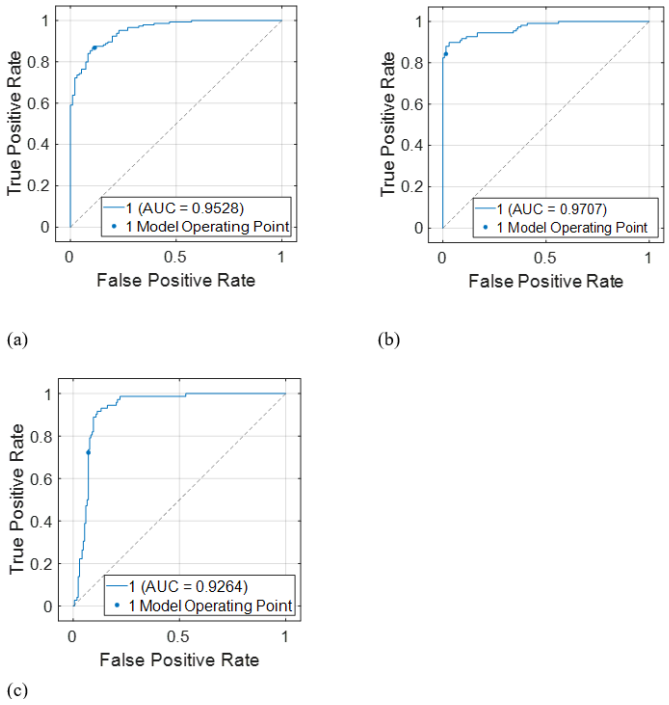


Figure 8. RUC curve of test data using SVM models: (a) trained and tested on EL48+H48 data, (b) trained and tested on EL36+H36 data, and (c) trained and tested on EL24+H24 data.

4. Conclusions

A portable spectrometer was used successfully to classify durian chips into two categories: chips fried in oil used for 48-hour duration or less and chips fried in oil used for more than that duration, which resulted in total polar compounds exceeding 25%. The fusion strategy that involved first combining all the preprocessed NIR spectral data and selecting feature wavelengths using GA had better performance with PLS-DA than from using either of the two individually preprocessed spectral

datasets. The best classification using the SVM models achieved a precision of 91.9%. The results highlighted the advantage of the alternative fusion technique, which in this study was implemented by combining preprocessing methods that used only NIR spectral data from a single instrument. However, the robustness of the SVM model was unsatisfactory when tested on different datasets. When evaluated using the same dataset used for model development, the best classification performance with a precision of 97.8% was achieved for distinguishing between chips fried in oil used for 36 h or less and those fried in oil used for more than 36 h.

Overall, when integrated with a mobile application for control and real-time display of classification results, the portable NIR spectrometer provided a rapid, on-site solution for evaluating the quality of durian chips fried in repeated-use oil. Thus, this approach is highly suitable for integration into real-time quality control processes during chip production. However, to maintain classification accuracy, it is essential to control factors, such as chip thickness, moisture content, and other physical or chemical variations. Furthermore, regular maintenance and routine calibration of the spectrometer are necessary to ensure consistent and reliable performance over time. Finally, the classification model's performance could be further enhanced by applying advanced machine learning techniques in the analysis stage.

Author contributions

Arthit Phuangsombut: Visualization, Investigation, Methodology; Kaewkarn Phuangsombut: Methodology, Data curation, Validation; Sirinad Noypitak: Data curation, Resources, Software; Anupun Terdwongworakul: Writing – original draft, Writing – review & editing, Supervision, Conceptualization, Formal analysis. All authors have read and agreed to the published version of the manuscript.

Use of Generative-AI tools declaration

The authors declare that they have not used artificial intelligence tools in the creation of this article.

Acknowledgments

The authors gratefully acknowledge research support from the Faculty of Engineering at Kamphaeng Saen, Kasetsart University, Thailand.

Conflict of interest

The authors declare no conflict of interest.

References

1. Sun Y, Zhang M, Fan D (2019) Effect of ultrasonic on deterioration of oil in microwave vacuum frying and prediction of frying oil quality based on low field nuclear magnetic resonance (LF-NMR). *Ultrason Sonochem* 51: 77–89. https://doi.org/10.1016/j.ultsonch.2018.10.015

2. Gertz C, Behmer D (2014) Application of FT-NIR spectroscopy in assessment of used frying fats and oils. *Eur J Lipid Sci Technol* 116: 756–762. https://doi.org/10.1002/ejlt.201300270
3. Hassanien MFR, Sharoba AM (2014) Rheological characteristics of vegetable oils as affected by deep frying of French fries. *Food Measure* 8: 171–179. https://doi.org/10.1007/s11694-014-9178-3
4. Feng H, Li Y, Sui X, et al. (2016) Effect of frying cycles on polar components in soybean oil and absorbed lipids of fried potatoes. *Trans Chin Soc Agric Eng* 32: 309–314. https://doi.org/10.11975/j.issn.1002-6819.2016.03.044
5. Sanchez-Muniz FJ, Cuesta C, Garrido-Polonio C (1993) Sunflower oil used for frying: combination of column, gas and high-performance size-exclusion chromatography for its evaluation. *JAOCS* 70: 235–240. https://doi.org/10.1007/BF02545301
6. Hosseini H, Ghorbani M, Meshginfar N, et al. (2016) A review on frying: procedure, fat, deterioration progress and health hazards. *JAOCS* 93: 445–466. https://doi.org/10.1007/s11746-016-2791-z
7. Ng CL, Wehling RL, Cuppett SL (2007) Method for determining frying oil degradation by near-infrared spectroscopy. *J Agri Food Chem* 55: 593–597. https://doi.org/10.1021/jf061841d
8. Ng CL, Wehling RL, Cuppett SL (2011) Near-infrared spectroscopic determination of degradation in vegetable oils used to fry various foods. *J Agri Food Chem* 59: 12286–12290. https://doi.org/10.1021/jf202740e
9. Öğütcü M, Aydeniz B, Büyükcan MB (2012) Determining frying oil degradation by near infrared spectroscopy using chemometric techniques. *JAOCS* 89: 1823–1830. https://doi.org/10.1007/s11746-012-2087-x
10. Kazemi S, Wang N, Ngadi M, et al. (2005) Evaluation of frying oil quality using VIS/NIR hyperspectral analysis. *Agric Eng Int: CIGR J* VII: EP05 001.
11. Liu Y, Sun L, Li Y, et al. (2018) Quality evaluation of fried soybean oil base on near infrared spectroscopy. *J Food Process Eng* 41: e12887. https://doi.org/10.1111/jfpe.12887
12. Liu Y, Sun L, Bai H, et al. (2020) Detection for frying times of various edible oils based on near-infrared spectroscopy. *Appl Sci* 10: 7789. https://doi.org/10.3390/app10217789
13. Liu Y, Sun L, Du C, et al. (2020) Near-infrared prediction of edible oil frying times based on Bayesian Ridge Regression. *Optik* 218: 164950. https://doi.org/10.1016/j.ijleo.2020.164950
14. Ma J, Zhang H, Tuchiya T, et al. (2014) Rapid determination of degradation of frying oil using near-infrared spectroscopy. *Food Sci Technol Res* 20: 217–223. https://doi.org/10.3136/fstr.20.217
15. García-Martín JF, Alés-Álvarez FJ, del Carmen López-Barrera M, et al. (2019) Cetane number prediction of waste cooking oil-derived biodiesel prior to transesterification reaction using near infrared spectroscopy. *Fuel* 240: 10–15. https://doi.org/10.1016/j.fuel.2018.11.142
16. Isdhiyanti MIT, Budiastra IW, Mala DM, et al. (2024) Prediction of chemical contents of cooking oil using near-infrared spectroscopy. *Food Res* 8: 147–152. https://doi.org/10.26656/fr.2017.8(6).520
17. Maniwara P, Meesombat R, Malang S, et al. (2024) Determination of oil quality during crispy pork rind frying: Near infrared spectra and color values as predictors. *J Food Eng* 383: 112251. https://doi.org/10.1016/j.jfoodeng.2024.112251

18. Palarea-Albaladejo J, Cayuela-Sánchez JA, Moriana-Correro E (2022) Estimating fat components of potato chips using visible and near-infrared spectroscopy and a compositional calibration model. *Food Anal Methods* 15: 133–143. https://doi.org/10.1007/s12161-021-02106-0
19. Pedreschi F, Segtnan VH, Knutsen SH (2010) On-line monitoring of fat, dry matter and acrylamide contents in potato chips using near infrared interactance and visual reflectance imaging. *Food Chem* 121: 616–620. https://doi.org/10.1016/j.foodchem.2009.12.075
20. Areekij S, Ritthiruangdej P, Kasemsumran S, et al. (2017) Rapid and nondestructive analysis of deep-fried taro chip qualities using near infrared spectroscopy. *J Near Infrared Spectrosc* 25: 127–137. https://doi.org/10.1177/0967033516686655
21. López Maestresalas A, López Molina C, Oliva Lobo GA, et al. (2022) Evaluation of near-infrared hyperspectral imaging for the assessment of potato processing aptitude. *Front Nutr* 9: 999877. https://doi.org/10.3389/fnut.2022.999877
22. Zhang Y, Wang Y (2023) Machine learning applications for multi-source data of edible crops: A review of current trends and future prospects. *Food Chem*: X, 100860. https://doi.org/10.1016/j.fochx.2023.100860
23. Di Rosa AR, Leone F, Cheli F, Chiofalo V (2017) Fusion of electronic nose, electronic tongue and computer vision for animal source food authentication and quality assessment–A review. *J Food Eng* 210: 62–75. https://doi.org/10.1016/j.jfoodeng.2017.04.024
24. Xia H, Chen W, Hu D, et al. (2024) Rapid discrimination of quality grade of black tea based on near-infrared spectroscopy (NIRS), electronic nose (E-nose) and data fusion. *Food Chem* 440: 138242. https://doi.org/10.1016/j.foodchem.2023.138242
25. Hong Y, Birse N, Quinn B, et al. (2023) Data fusion and multivariate analysis for food authenticity analysis. *Nat Commun* 14: 3309. https://doi.org/10.1038/s41467-023-38382-z
26. Xu P, Fu L, Xu K, et al. (2023) Investigation into maize seed disease identification based on deep learning and multi-source spectral information fusion techniques. *J Food Compos Anal* 119: 105254. https://doi.org/10.1016/j.jfca.2023.105254
27. An H, Zhai C, Zhang F, et al. (2023) Quantitative analysis of Chinese steamed bread staling using NIR, MIR, and Raman spectral data fusion. *Food Chem* 405:134821.
28. Cevoli C, Iaccheri E, Fabbri A, et al. (2024) Data fusion of FT-NIR spectroscopy and Vis/NIR hyperspectral imaging to predict quality parameters of yellow flesh "Jintao" kiwifruit. *Biosyst Eng* 237: 157–169. https://doi.org/10.1016/j.biosystemseng.2023.12.011
29. Ping J, Ying Z, Hao N, et al. (2024) Rapid and non-destructive identification of Panax ginseng origins using hyperspectral imaging, visible light imaging, and X-ray imaging combined with multi-source data fusion strategies. *Int Food Res* 192: 114758. https://doi.org/10.1016/j.foodres.2024.114758
30. Li X, Cai M, Li M, et al. (2023) Combining Vis-NIR and NIR hyperspectral imaging techniques with a data fusion strategy for the rapid qualitative evaluation of multiple qualities in chicken. *Food Control* 145: 109416. https://doi.org/10.1016/j.foodcont.2022.109416
31. Feng Y, Dong F, Chen Y, et al. (2024) Combining Vis-NIR and NIR hyperspectral imaging techniques with a data fusion strategy for prediction of norfloxacin residues in mutton. *Spectrochim Acta A Mol Biomol Spectrosc* 322: 124844. https://doi.org/10.1016/j.saa.2024.124844

32. Posom J, Phuphaphud A, Saengprachatanarug K, et al. (2022) Real-time measuring energy characteristics of cane bagasse using NIR spectroscopy. *Sens BioSensing Res* 38: 100519. https://doi.org/10.1016/j.sbsr.2022.100519
33. Leardi R (2000) Application of genetic algorithm–PLS for feature selection in spectral data sets. *J Chemom* 14: 643–655. https://doi.org/10.1002/1099-128X(200009/12)14:5/6%3C643::AID-CEM621%3E3.0.CO;2-E
34. Haupt, SE (2009) Introduction to genetic algorithms, In: *Artificial Intelligence Methods in the Environmental Sciences*, 103–125, Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-9119-3_5
35. Ali MM, Hashim, N, Shahamshah, MI (2021) Durian (Durio zibethinus) ripeness detection using thermal imaging with multivariate analysis. *Postharvest Biol Technol* 176: 111517. https://doi.org/10.1016/j.postharvbio.2021.111517
36. Ditcharoen S, Sirisomboon P, Saengprachatanarug K, et al. (2023) Improving the non-destructive maturity classification model for durian fruit using near-infrared spectroscopy. *Artif Intell Agric* 7: 35–43. https://doi.org/10.1016/j.aiia.2023.02.002
37. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14: 199–222. https://doi.org/10.1023/B:ASTCO.0000035301.49549.88
38. Raschka S, Mirjalili V (2019) Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2nd Edition, Birmingham: Packt Publishing.
39. Osborne BG, Fearn T, Hindle PH (1993) Practical NIR Spectroscopy with Application in Food and BeVerage Analysis, 2 Eds., Essex: Longman Scientific & Technical, 13–35.
40. Workman J, Weyer L (2012) Practical Guide and Spectral Atlas for Interpretive Near Infrared Spectroscopy, 2 Eds., Boca Raton: CRC Press. https://doi.org/10.1201/b11894

AIMS Press