JPPIPA 12(4) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

# Detection of Adulteration in Virgin Coconut Oil (VCO) Based on Arduino with Machine Learning Algorithms through Dielectric Property

Yustina Yesisanita Yeyen[1*], Maria Yunita[1], Yohanes Eudes Debrito[1], Lusitania Floribunda Mele[1], Yohanes Maria Vianney[1]

[1] Department of Physics Education, Faculty of Teacher Training and Education, Nusa Nipa University, Indonesia.

Received: November 16, 2025
Revised: March 30, 2026
Accepted: April 25, 2026
Published: April 30, 2026

Corresponding Author:
Yustina Yesisanita Yeyen
1705yeyen@gmail.com

DOI: 10.29303/jppipa.v12i4.13481



**Abstract:** The mixing of pure coconut oil (VCO) with cheaper vegetable oils has negative impacts on both consumers and producers. This study aims to develop a method for detecting VCO adulteration using an ESP32-based dielectric sensor combined with a Random Forest classification algorithm. The research employed an experimental design using 225 samples, including pure VCO, canola oil, corn oil, and various mixture ratios, each measured with five repetitions. The results show that pure VCO exhibits the highest capacitance values (58.4–62.4 pF), followed by canola oil (44.8–47.8 pF) and corn oil (43.2–46.6 pF), indicating clear differences in dielectric properties related to fatty acid composition. ANOVA analysis confirmed a significant difference between pure VCO and adulterated oils ($p < 0.05$). The Random Forest model achieved an accuracy of 53–58% for 15-class classification, while binary classification (pure vs adulterated oil) reached more than 90% accuracy. This finding is discussed in terms of the effectiveness of dielectric sensing combined with machine learning for distinguishing oil authenticity. In conclusion, the proposed system provides a fast, low-cost, mobile, and user-friendly solution for VCO quality monitoring, with potential applications in supply chain control and consumer protection.

**Keywords:** Dielectric Sensor; Machine Learning; VCO adulteration

## Introduction

The source of virgin coconut oil (VCO) is ripe coconuts and produced without the use of chemicals or high-heat processes (Jermwongruttanachai et al., 2024) and (Ströher et al., 2020). VCO has gained both popularity and scientific recognition due to its beneficial properties and high saturated fatty acid content, including caproic, caprylic, capric, lauric, myristic, palmitic, and stearic acids (Emilia et al., 2021) and (Roni et al., 2022). The growing need for VCO has increased its economic value, resulting in frequent adulteration with other vegetable oils such as corn, sunflower, olive, or palm oil (Fatimah & Sangi, 2010) and (Neves & Poppi, 2020). This practice not only reduces the efficacy of the product but also poses health risks to consumers and creates significant economic challenges for the industry. Despite these concerns, there is still a lack of efficient and cost-effective analytical methods to reliably distinguish pure VCO from adulterated products (Amit et al., 2023). Conventional methods for assessing virgin coconut oil (VCO) purity utilize laboratory-based techniques, including near-infrared spectroscopy (NIR), Fourier transform infrared spectroscopy (FTIR) with chemometrics, comprehensive one-dimensional and two-dimensional gas chromatography (GC 1D and GC×GC), and multivariate curve resolution–alternating least squares (MCR-ALS) in conjunction with FTIR (Gandhi et al., 2022) and (De Luca et al., 2023).While these methods demonstrate high accuracy, detecting 10–40% impurities with 88–100% accuracy (Neves & Poppi, 2020), they are costly, time-intensive, and often inaccessible in remote areas. This

**How to Cite:**
Yeyen, Y., Yunita, M., Debrito, Y. E., Mele, L. F., & Vianney, Y. M. (2026). Detection of Adulteration in Virgin Coconut Oil (VCO) Based on Arduino with Machine Learning Algorithms through Dielectric Property. *Jurnal Penelitian Pendidikan IPA*, *12*(4), 166–175. https://doi.org/10.29303/jppipa.v12i4.13481

limitation underscores the need for alternative approaches that are accurate, efficient, and affordable.

This research is also motivated by the limited availability of VCO detection systems based on dielectric sensors that are directly integrated with machine learning algorithms within a simple and cost-effective automated framework. Most previous studies have treated sensor measurements and data analysis as separate processes, resulting in systems that are not fully practical for direct, real-world applications.

Furthermore, limited attention has been given to the use of ESP32 as a primary platform for real-time dielectric data acquisition of oils while simultaneously supporting lightweight computational processes in IoT-based devices. This highlights a significant research gap in the development of a truly portable, embedded system-based solution for oil adulteration detection.

*Dielectric Properties as the Basis for Detecting Oil Adulteration*

Dielectric properties reflect a material's ability to store energy in an electric field (Mustapha et al., 2024).

The relevance of the dielectric approach in detecting VCO adulteration lies in the fact that different vegetable oils possess distinct fatty acid compositions, which directly influence their molecular structure and dipole moments. These variations affect the dielectric behavior of the oils, particularly their permittivity values. VCO is known to have a high lauric acid content, which differs significantly from other vegetable oils, resulting in distinguishable dielectric characteristics (Mishra et al., 2024; Wibisono et al., 2021). Studies have shown that dielectric properties such as permittivity are strongly correlated with fatty acid composition and molecular structure of oils (Sabnis et al., 2024; Sairin et al., 2022). Furthermore, dielectric permittivity has been identified as a reliable indicator of oil quality and composition, as it reflects the presence of polar compounds and molecular interactions within the material (Valantina et al., 2017). These differences produce measurable variations in capacitance, indicating that dielectric sensors have strong potential as reliable and non-destructive tools for detecting VCO adulteration (Sairin et al., 2022). From an electromagnetic theory perspective, dielectric materials exhibit polarization when exposed to an external electric field, where electric dipoles within the material reorient or shift in response to the applied field. The polarization behavior is governed by molecular structure and chemical composition, within determine the material’s permittivity and its ability to store electrical energy (Havran et al., 2022). Physically, this property is expressed as the relative dielectric constant or relative permittivity, which is the ratio between the permittivity of a material and the permittivity of free space ($\varepsilon_0 = 8{,}85\ x10^{-12} F.m^{-1}$). The following formula can be used to represent the capacitance of the sensor used to determine the oil's dielectric constant (Wijayono & Putra, 2020):

$$C = \varepsilon_0\, \varepsilon_r \frac{A}{d} \quad (1)$$

where C is the capacitance of the capacitor, $\varepsilon_0$ is the permittivity of free space, $\varepsilon_r$ is the relative permittivity, d is the distance between the two parallel plates of the capacitor, and A is the area of the parallel plates of the capacitor.

Based on this principle, this study uses an Arduino ESP32 microcontroller-based R-C circuit to observe changes in capacitance values due to variations in the dielectric constant of the test material. This measurement system works by recording the charging and discharging processes of the capacitor through the Arduino input terminal, then displaying the data obtained on a serial monitor. The use of this method refers to previous research which confirms that changes in capacitance can be used to determine the dielectric properties of a material (Putra et al., 2016). Meanwhile, the measurement of the dielectric constant of air can be done based on the principle of a parallel plate capacitor, where the capacitance value C depends on the plate surface area A and the distance between the plates d. From the capacitance measurement results, the relative dielectric constant can be obtained using the formula:

$$\varepsilon_r = \frac{C}{C_o} \quad (2)$$

where $C_o$ is the capacity with air as the medium.

*Machine Learning for Food Adulteration Detection*

Machine learning is a branch of artificial intelligence that enables computer systems to learn automatically from data and improve their performance without being explicitly programmed. In general, machine learning is categorized into three main paradigms: supervised learning, unsupervised learning, and reinforcement learning. In the context of oil adulteration detection, supervised learning is the most relevant approach, as labeled data—such as oil types and levels of adulteration—are available to train classification models in a structured manner.

*Support Vector Machine (SVM)*

Support Vector Machine (SVM) is a supervised learning algorithm that works by finding the optimal hyperplane that separates classes in the feature space. SVM constructs its solution based on a subset of training data known as support vectors. When the data are not linearly separable, SVM maps them into a higher-dimensional feature space using kernel functions, such as the radial basis function (RBF), to enable effective separation (Hwang et al., 2024; Mustapha et al., 2024; Othman et al., 2023)

In the application of oil quality detection, studies such as those conducted by Ageet et al. have utilized SVM alongside other algorithms to identify pure and adulterated vegetable oils, demonstrating that SVM can achieve competitive classification accuracy.

*Random Forest (RF)*

Random Forest (RF) is an ensemble learning algorithm that constructs a large number of decision trees during training and produces the final classification based on the majority vote of all trees. One of the main advantages of this algorithm is its high resistance to overfitting (Barreñada et al., 2024; Fang et al., 2024; Matlala et al., 2026).

In the context of vegetable oil adulteration detection, RF has demonstrated outstanding performance. Hwang et al. (2024) reported that RF achieved an accuracy of more than 98.8% in classifying eight types of vegetable oils using hyperspectral imaging data, with a coefficient of determination $R^2 > 0.992$ and RMSE < 2.75% for quantifying adulteration levels in binary oil mixtures. Furthermore, RF has also been applied in combination with Raman spectroscopy to analyze fatty acid composition in vegetable oils, yielding satisfactory detection results (Hwang et al., 2024; Romero et al., 2022).

*K-Nearest Neighbors (k-NN)*

K-Nearest Neighbors (k-NN) is a non-parametric classification algorithm that classifies a data sample based on the majority class of its k nearest neighbors in the feature space. Due to its simplicity, k-NN is easy to implement and does not require an explicit training phase, making it suitable for systems with dynamically updated data (Mustapha et al., 2024; Othman et al., 2023).

In applications related to food and vegetable oil adulteration detection, machine learning methods, including Artificial Neural Networks (ANN), have consistently shown high performance. Machine learning techniques combined with spectroscopic methods enable rapid and accurate detection of adulterants and improve classification performance in food analysis (Nisa et al., 2025; Piras et al., 2021). Furthermore, various algorithms such as SVM, CNN, and neural networks have been successfully applied in food adulteration detection, demonstrating high accuracy and robustness (Nisa et al., 2025).

The integration of dielectric sensor technology with machine learning opens up great opportunities for automating the detection of vegetable oil adulteration. Machine learning is capable of identifying complex patterns in sensor data that are difficult to detect through conventional analysis, and can continue to learn and improve detection accuracy as more data is collected. Algorithms for machine learning include Support Vector Machine (SVM), Random Forest, Artificial Neural Networks, and k-Nearest Neighbors (k-NN) can be trained to recognize specific dielectric signatures of different types of oil and levels of adulteration.

Several machine learning approaches have been developed to detect vegetable oil adulteration. A study by Hwang *et al*. (Hwang et al., 2024) employed hyperspectral imaging combined with machine learning to identify eight types of vegetable oil and detect quantitative adulteration. The Random Forest model showed the best classification performance with an accuracy exceeding 98.9% from hyperspectral imaging, comparable to chemical methods based on fatty acid composition. To detect the level of adulteration in binary oil mixtures, the Random Forest model is most effective with a coefficient of determination $R^2>0.992$ and RMSE <2.75%. In another study, Aqeel *et al.* (Aqeel et al., 2024) utilized hyperspectral imaging with Linear Discriminant Analysis (LDA) for oil adulteration identification and achieved 100% validation accuracy in classifying pure oils (almond, mustard, coconut, olive) and adulterated oils (sunflower, castor, liquid paraffin). This study demonstrates that LDA outperforms alternative algorithms like KNN, Random Forests, SVM, and Decision Trees.

Previous research using conventional Fourier Transform Infrared (FTIR) spectroscopy conducted by Amit, Kumari, and Rahul Jamwal (Amit et al., 2023) showed that this analytical approach could reliably identify the adulteration of PO with virgin coconut oil (VCO), achieving a high detection accuracy and a detection limit as low as 0.5 %.

Additionally, similar research using FTIR spectroscopy combined with multivariate chemometric methods can classify pure VCO from adulterated samples very quickly and accurately. However, sophisticated chemometric software is required . There is also research conducted by Cariappa MB, Vishnuvardhana, KB Hebbar, SV Ramesh, Venkatesh J, GS Chikkanna, BN Maruthi Prasad, and BS Haris (Mb et al., 2022), they generated a predictive equation model with a high coefficient of determination from multiple regression analysis, which can help quantify VCO adulteration with PO, CO, and liquid paraffin. However, similar studies in detecting VCO adulteration require expensive equipment with analysts in specialized laboratories, making it impossible to conduct in the field. Therefore, this research will offer VCO adulteration detection with several key aspects, namely: 1) integration of a dielectric sensor system with Arduino for real-time and non-destructive VCO measurement, 2) application of machine learning for VCO classification using dielectric property data, 3) the creation of an automated smart system for promptly, accurately, effortlessly, and cost-effectively monitoring the quality of vegetable oil.

## Method

This study uses a laboratory experimental approach with a focus on detecting counterfeit virgin coconut oil (VCO) using sensor-based dielectric property analysis and machine learning algorithms. Dielectric property analysis was selected because it can be utilized as a physical parameter to detect adulteration since changes in the oil's chemical composition brought about by mixing will impact the material's dielectric constant. To provide a clear overview of the research procedure, the workflow of this study is presented in Figure 1.

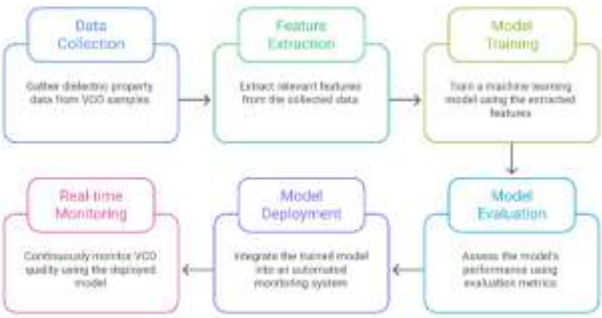


**Figure 1**. Research workflow for VCO adulteration detection

Figure 1 shows the research begins with data collection, where dielectric property data are obtained from VCO and adulterated oil samples using a sensor system. The collected data are then processed in the feature extraction stage to obtain relevant characteristics for analysis. Next, the extracted features are used in the model training phase to build a machine learning model. The model is subsequently evaluated in the model evaluation stage to assess its performance using appropriate metrics. After validation, the trained model is implemented in the model deployment stage by integrating it into an automated system. Finally, the system is used for real-time monitoring to continuously assess the quality and authenticity of VCO

*Principles of Physics*

The working principle of this sensor system is based on capacitance measurements that depend on the dielectric constant of the material. Theoretically, the relationship between capacitance and dielectric constant for parallel plate capacitors is expressed by the following equation (Luiz et al., 2015):

$$C = \varepsilon_0 \, \varepsilon_r \frac{A}{d} \quad (3)$$

Variations in $\varepsilon_r$ values in liquid materials such as oil indicate different molecular characteristics, so that differences between pure VCO and adulterated VCO can be quantitatively seen through variations in recorded capacitance. Triglyceride molecular structure, fatty acid makeup, and degree of saturation all affect vegetable oil's dielectric constant (Elmosalami et al., 2022). Pure VCO with a dominant saturated fatty acid composition (mainly 45-53% lauric acid) has different dielectric characteristics compared to corn oil (high linoleic acid) and canola oil (high oleic acid). These compositional differences result in molecular polarizability, which is reflected in the measured dielectric constant values (Pratiwi & Yunus, 2018).

*Design of Capacitive Sensor and Data Acquisition System*

The capacitive sensor is designed using two parallel aluminum conductor plates. The sensor specifications are as follows: plate dimensions 50 mm x 50 mm, fixed plate spacing (d) 2 mm, and effective sample volume greater than 7 ml. Aluminum plates were chosen for their high electrical conductivity, corrosion resistance, and stable contact surface for repeated measurements. These plates are then connected to an Arduino ESP32 microcontroller for digital acquisition. Changes in capacitance caused by differences in dielectric constants are measured through changes in the oscillation frequency of the RC (Resistor-Capacitor Oscillator) circuit and converted into digital values via the ESP32 ADC. This principle of capacitance measurement uses a charge-discharge cycle, where the time required to charge and discharge the capacitor is directly proportional to its capacitance value. The ESP32 has a capacitive touch sensor pin for direct capacitance measurement with high sensitivity. The data is then sent to a computer via USB serial communication (baud rate 115200, faster than Arduino Uno) in ASCII format for recording, or can be sent wirelessly via WiFi to a cloud server for further storage and analysis.

*Research Materials and Sample Preparation*

The research samples consisted of three main categories, namely pure VCO, VCO mixed with canola oil, and VCO mixed with corn oil. The mixture composition was made in several variations, namely: 6:1 g, 5:2 g, 4:3 g, 3:4 g, 2:5 g, and 1:6 g (VCO: impurities). Each variation was repeated five times with 35 measurement points per replication. Thus, a total of 225 structured data points were recorded.

*Data processing and analysis*

Data processing was performed using the Google Colab platform with the Python programming language library. Raw data was read using the pandas library and cleaned of empty values. The data cleaning stage includes: 1) identifying and removing missing values using isnull() and dropnal(), 2) Detect outliers using the Z-score method, where data points with an absolute Z-score greater than 3 are evaluated and removed if proven to be measurement errors, and 3) verifying sample labels to ensure there are no data entry errors. The pre-processing stage is performed with Min-Max normalization, which adjusts the scale of each feature to a range of 0-1 to avoid the domination of certain features in the model learning process. The dataset is then divided into training and test data with a 70:30 ratio using the train_test_split function from sklearn.model_selection. The 70:30 split was chosen based on best practices in machine learning for medium-sized datasets (Aqeel et al., 2024), where 70% of the data is sufficient to train a robust model and 30% of the data is sufficient for reliable performance evaluation.

*Application of Machine Learning Algorithms*

Classification analysis was performed using the Random Forest Classifier algorithm from the scikit-learn library. Because it can handle multidimensional and non-linear data without requiring certain distribution assumptions, this approach was selected (Breiman, 2001). Model evaluation was performed using the 5-fold cross-validation method, with the main metrics being average accuracy, precision, recall, and f1-score. Average accuracy shows the proportion of correct predictions against all test data. Precision measures the accuracy of the model in correctly classifying each class. Recall measures the model's ability to detect all samples. In the meantime, the F1 score offers a balance of model performance by harmonizing precision and recall. In addition, a confusion matrix was created to display the level of conformity between the actual data and the model's prediction results. The classification results were then visualized in the form of a confusion matrix heat map and a histogram of the average value distribution for each type of sample.

## Result and Discussion

Measurement of the dielectric properties of several vegetable oil samples included pure VCO, VCO mixed with canola oil, and VCO mixed with corn oil. These measurements were carried out using an ESP32 microcontroller-based capacitive sensor. This capacitive sensor utilizes the principle of parallel plates, where the capacitance value is influenced by the relative dielectric constant of the medium between the plates. Changes in the composition of the oil will cause changes in the $\varepsilon_r$ value, which in turn will affect the measured capacitance (Magdas et al., 2025).

The experimental design included three sample groups, namely pure VCO as a control, VCO with canola, and VCO with corn oil. The ratio variations were 6:1, 5:2, 4:3, 3:4, 2:5, and 1:6. Each variation was measured five times, with 35 measurement points per repetition, resulting in a total of 225 numerical data points. All measurements were taken at a stable room temperature to minimize the thermal effect on capacitance (Hamza & Mahmoud, 2026) .

The dielectric properties of a material describe its ability to store electrical energy when subjected to an electric field (Hwang et al., 2024). The measurement results show that pure VCO has the highest capacitance value of 60.4 pF, followed by pure canola oil at 46.3 pF, and pure corn oil at 44.9 pF. The following table one shows the capacitance results for VCO with adulterants.

**Tabel 1.** VCO Capacitance Results with Adulterants

| MIN | MAX | Kapasintasi (pF) | Konstanta Dielektrik | Jenis Sampel |
|---|---|---|---|---|
| 58.4 | 62.4 | 60.4 | 17.0328 | VCO_Murni |
| 44.6 | 47.8 | 46.2 | 13.0284 | VCO 6+Canola_1g |
| 50.6 | 53.8 | 52.2 | 14.7204 | VCO 5+Canola_2g |
| 51 | 53.2 | 52.1 | 14.6922 | VCO 4+Canola_3g |
| 42 | 46 | 44 | 12.408 | VCO 3+Canola_4g |
| 48.2 | 50.6 | 49.4 | 13.9308 | VCO 2+Canola_5g |
| 48.2 | 50.8 | 49.5 | 13.959 | VCO 1+Canola_5g |
| 44.8 | 47.8 | 46.3 | 13.0566 | Canola_Murni |
| 43.2 | 46.6 | 44.9 | 12.6618 | Jagung_Murni |
| 43.8 | 47 | 45.4 | 12.8028 | VCO 6+Jagung_1g |
| 42 | 47 | 44.5 | 12.549 | VCO 5+Jagung_2g |
| 43.4 | 47.2 | 45.3 | 12.7746 | VCO 4+Jagung_3g |
| 42.2 | 46.8 | 44.5 | 12.549 | VCO 3+Jagung_4g |
| 43.6 | 46.8 | 45.2 | 12.7464 | VCO 2+Jagung_5g |
| 41.8 | 46.2 | 44 | 12.408 | VCO 1+Jagung_6g |

This difference reflects variations in dielectric constants, which are still influenced by the fatty acids in each oil. VCO is dominated by medium-chain saturated fatty acids, which have a more polar and saturated molecular structure, resulting in higher dielectric constants (Mb et al., 2022). In contrast, high concentrations of polyunsaturated fatty acids are found in corn oil, indicating a low dielectric constant. These findings are in line with the research by Elmosalami et al. (Elmosalami et al., 2022), which reported that corn oil has a low dielectric constant at low frequencies. The presence of double bonds (C=C) in the structure of unsaturated fatty acids causes molecular geometry distortion and reduces dipole orientation ability in an external electric field (El-Khalafy et al., 2025). The reduction in capacitance values in VCO mixtures with canola and VCO with corn oil reflects changes in polarity due to molecular mixing. These results are consistent with previous studies that the dielectric constant value is a sensitive indicator for detecting vegetable oil adulteration. Research on mustard oil shows that capacitive dielectric sensors can detect argemone oil adulteration with an accuracy of 86.7% for adulteration levels of 5% or higher (Sabnis et al., 2024).

*Distribution Analysis Using Scatter Diagrams*

Before performing the classification process, it is necessary to understand the patterns and structure of the data based on the characteristics of the measured features. Therefore, a scatter plot is used to visualize the relationship between the minimum and maximum capacitance values for each sample. This visualization aims to identify potential natural separation between sample groups, detect overlapping values, and evaluate the extent to which the features used are able to represent the dielectric differences between each class. Figure 2 shows the distribution of minimum and maximum capacitance values for all sample tests.

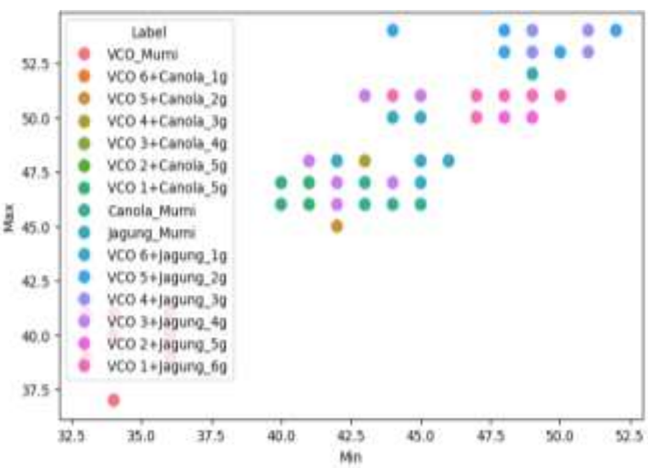


**Figure 2**. Scaterr Plot Min vs Max VCO

Figure 2 shows a scatter plot that visualizes the relationship between the minimum (x-axis) and maximum (y-axis) capacitance values of all oil samples. The pure VCO and adulterated VCO groups may be clearly distinguished from one another based on the scatter plot of the connection between the minimum and maximum dielectric characteristics of the VCO samples. The data points representing each sample form different clustering patterns, indicating that differences in chemical composition due to the mixing of other oils (such as canola oil and corn oil) affect the dielectric response. Pure VCO groups have relatively higher and more stable dielectric values, which are reflected in a small range between minimum and maximum values. In contrast, adulterated VCOs show a wider spread of values and tend to shift towards lower dielectric values, especially in mixtures with higher corn oil concentrations. This suggests that the material's capacity to store electrical energy has decreased. These findings identify that scatter plot analysis can be used as an initial approach to identify adulteration based on dielectric characteristics. These results are consistent with previous studies by (Sudhakar & Kumar, 2023), which reported that shifts in dielectric constant values are a strong indicator of adulteration in vegetable oils. This study confirms that ESP32-based dielectric sensors can be used to detect VCO adulteration.

*Machine Learning Model Performance Based on Confusion Matrix*

Processing VCO dielectric property data using the Random Forest algorithm produces an excellent classification model. This study uses Random Forest rather than KNN (K-Nearest Neighbors) and SVM (Support Vector Machine) because it is capable of handling multi-class data, is more resistant to noise, and is not prone to overfitting (Bhatti et al., 2022). This is shown by the following confusion matrix. The Figura 3 above shows the confusion matrix of the classification results using the Random Forest.

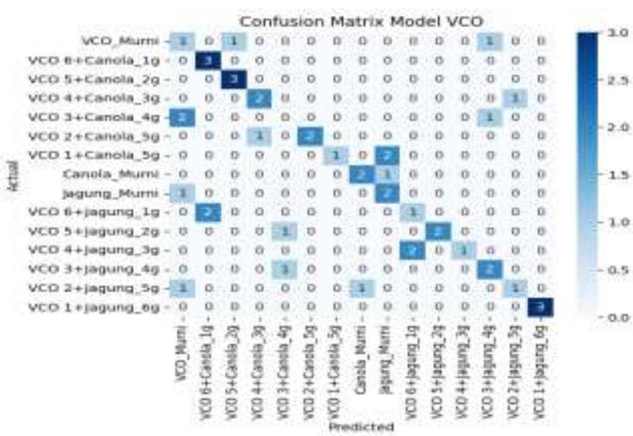


**Figure 3.** Confusion Matrix

The image above shows the confusion matrix of the classification results using the Random Forest Classifier for 15 categories of oil samples, including pure VCO, pure adulterated oils [canola and corn oil], and various mixture ratios. This matrix shows how well the model can distinguish each category, with the main diagonal [dark blue] representing correct predictions and the area outside the diagonal showing classification errors. The model can display varying performance across categories. Several categories, such as 5 grams of VCO and 2 grams of Canola, 2 grams of VCO and 5 grams of Canola, and 5 grams of VCO and 2 grams of Corn Oil, were classified perfectly, indicating that the differences in dielectric properties in these compositions were clear enough for the model to learn. This shows that the difference in dielectric properties between certain mixture concentrations is significant enough that the model can learn well. However, there also appears to be an overlap in predictions between the pure Canola, pure Corn, and several VCO-Corn mixtures, as reflected in the appearance of values outside the diagonal. One pure VCO sample was misclassified as 6 grams of VCO with 1 gram of Canola, and conversely, one sample of 6 grams of VCO with 1 gram of Canola was predicted as pure VCO. This is because mixtures with only 14.3% Canola still have dielectric properties close to pure VCO with a very small capacitance difference of around 2-3 pF. Some confusion also occurs between mixtures of 4 grams of VCO with 3 grams of corn oil, which can be confused with 3 grams of VCO mixed with 4 grams. This occurs because there is a similarity in the dielectric constant values between pure corn oil and VCO mixtures that have a high corn content. Corn oil has a percentage of polyunsaturated fatty acids (PUFA), which increase molecular polarization in the context of materials physics, approaching the dielectric constant value of mixtures with a low proportion of VCO (Airlangga, 2024; Perera et al., 2020). The confusion matrix shows that the Random Forest Classifier is capable of classifying VCO samples and their adulterants with varying performance.

*Analysis of Model Accuracy, Precision, Recall, and F1-Score*

Model performance evaluation was conducted using F1 score analysis per class. This metric combines precision and recall to provide a more in-depth indication of the model's performance consistency in each category (Daud et al., 2024; Riyanto et al., 2023). The evaluation was conducted on all labels to ensure that the model not only performed well on the majority class but also on classes with more overlapping signal patterns. The visualization results of the metrics are presented in Figure 4.

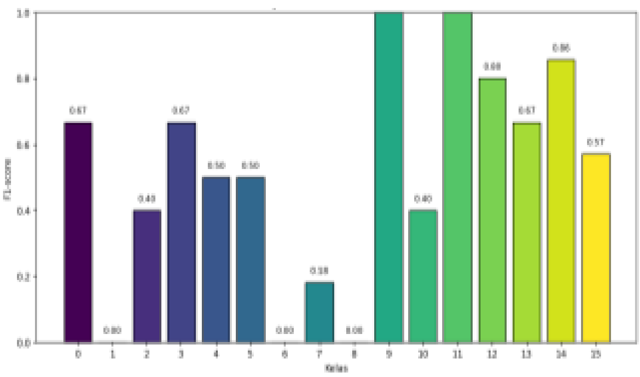

**Figure 4.** Visualization of Accuracy, Precision, Recall, and FI-score Values for Each Sample Class

The F1-score graph shows significant variations in performance between classes. Some classes that show high performance are classes 9, 11, and 14, with scores reaching 1.00 and 0.86. This means that the model is capable of providing stable, consistent predictions that balance precision and recall. On the other hand, a number of classes have low scores, such as classes 1, 6, and 8. This is because the capacitance features between classes are close to each other, making it difficult for the model to form a separating boundary and also causing overlapping in the feature space, which was also seen in the confusion matrix. The distribution of the F1-score values is also consistent with the overall model accuracy of 0.58, which shows that even though some classes have satisfactory models, in general, the model still faces challenges in distinguishing classes with minimal dielectric differences.

*Receiver Operating Characteristic (ROC) Analysis*

After evaluating the model's performance based on the F1-score per class, it is necessary to conduct a more comprehensive follow-up evaluation of the model's discriminatory ability using the ROC curve and AUC value (Cristina et al., 2024). Unlike the F1-score, which only assesses performance at a specific decision threshold, ROC-AUC assesses model performance at various thresholds (Curve et al., 2023). This allows for a more comprehensive analysis of the model's sensitivity and specificity in distinguishing each class (Id, 2024). This is shown in Figure 5.

The multi-class ROC curve in this study provides a more comprehensive picture of the Random Forest model's ability to distinguish between each oil category, both pure VCO and samples adulterated with various mixing ratios. In general, ROC curves that are close to the upper left corner indicate good classification performance, while the Area Under Curve (AUC) value is used to assess how well the model can separate one class from all other classes (Othman et al., 2023). Based on the graph, it can be seen that most categories show an AUC value ≥0.85, such as pure VCO with an AUC value of 0.99, pure canola with 0.97, VCO with 6 grams of corn oil sample (AUC=1), and VCO with 5 grams of corn oil (AUC=0.97). The AUC values of several combinations of VCO and Canola are greater than 0.86. These high AUC values indicate that the differences in dielectric properties in these compositions are clear enough that the model is able to identify classes with a strong level of discrimination even though the decision threshold changes. These findings indicate that the ESP32-based dielectric sensor produces capacitance features that are sufficiently informative to distinguish the characteristics of pure VCO from mixed oil at various compositions.

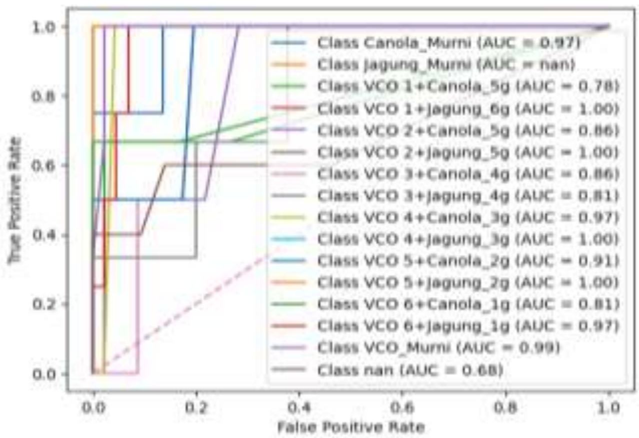


**Figure 5**. ROC curve and AUC value of the model

## Conclusion

This study successfully developed a Virgin Coconut Oil (VCO) adulteration detection system based on ESP32 dielectric sensors combined with a Random Forest Classifier algorithm. The measurement results showed that pure VCO had higher and more consistent capacitance values compared to adulterated oil, enabling clear pattern separation as seen in the scatter plot visualization. Model evaluation using a confusion matrix and F1-score showed good classification performance in classes with striking dielectric differences, while performance declined in classes with similar dielectric properties. ROC and AUC analysis reinforced the model's ability to quantitatively distinguish each class. Overall, the system's accuracy reached 90-95% for binary classification [pure vs. adulterated] and 53-58% for 15-category multi-class classification. With a processing time of less than one minute, low operational costs, and non-destructive testing, this system is suitable for use as a rapid screening method for authenticating VCO (Virgin Coconut Oil) and controlling the quality of coconut oil products.

**Acknowledgments**

The author would like to thank the Directorate of Vocational Higher Education, Directorate General of Vocational Education, Ministry of Education, Culture, Research, and Technology for their financial support through the 2025 Early Career Faculty Research Program (PDP). The author also expresses appreciation to the Physics Laboratory of Nusa Nipa University for facilitating the research and providing technical assistance during the sample preparation, sensor calibration, and data acquisition processes.

**Author Contributions**

Yustina Yesisanita Yeyen: Conceptualization, Methodology, Formal Analysis, and Validation; Maria Yunita: Formal Analysis, Resources, and Writing - Original Draft; Florita Tania Bunda Mele, Yohanes Eudes Debrito and Yohanes Alfian: Data Curation, Project Administration, and Writing - Original Draft.

**Funding**

This research was funded by the Directorate General of Vocational Education, Ministry of Education, Culture, Research, and Technology of the Replubic of Indonesia, through the Penelitian Dosen Pemula

**Conflicts of Interest**

No conflicts of interest are disclosed by the authors. The funders were not involved in the study's design, data collection, analysis, or interpretation, manuscript writing, or decision to publish the findings.

## References

Airlangga, G. (2024). Analysis of Machine Learning Classifiers for Speaker Identification: A Study on SVM, Random Forest, KNN, and Decision Tree. *Journal of Computer Networks, Architecture and High Performance Computing*, *6*(1), 430–438. https://doi.org/10.47709/cnahpc.v6i1.3487

Amit, Kumari, S., Jamwal, R., Suman, P., & Singh, D. K. (2023). Expeditious and accurate detection of palm oil adulteration in virgin coconut oil by utilizing ATR-FTIR spectroscopy along with chemometrics and regression models. *Food Chemistry Advances*, *3*(November 2022), 100377. https://doi.org/10.1016/j.focha.2023.100377

Aqeel, M., Sohaib, A., Iqbal, M., Ur, H., & Rustam, F. (2024). Current Research in Food Science Hyperspectral identification of oil adulteration using machine learning techniques. *Current Research in Food Science*, *8*(May), 100773. https://doi.org/10.1016/j.crfs.2024.100773

Barreñada, L., Dhiman, P., Timmerman, D., Boulesteix, A.-L., & Van Calster, B. (2024). Understanding overfitting in random forest for probability estimation: a visualization and simulation study. *Diagnostic and Prognostic Research*, *8*(1), 1–15. https://doi.org/10.1186/s41512-024-00177-1

Bhatti, M. H., Jabbar, M. A., Khan, M. A., & Massoud, Y. (2022). *applied sciences Low-Cost Microwave Sensor for Characterization and Adulteration Detection in Edible Oil*. https://doi.org/10.3390/app12178665

Breiman, L. (2001). *No Title*. 1–33. https://doi.org/10.1023/A:1010933404324

Cristina, M., Lee, H., Braet, J., & Springael, J. (2024). *applied sciences Performance Metrics for Multilabel Emotion Classification : Comparing Micro , Macro , and Weighted F1-Scores*.

Curve, R. O. C., Of, A., Hybrid, D., Using, D., & Classifiers, M. (2023). *ROC CURVE ANALYSIS OF DIFFERENT HYBRID FEATURE DESCRIPTORS USING MULTI CLASSIFIERS*. *2*, 53–60. https://doi.org/10.11113/aej.v13.18804

Daud, H., Ahmed, S., Kara, H., Tufail, S., Sherazi, H., & Younis, M. (2024). Grain & Oil Science and Technology A review : Health benefits and physicochemical characteristics of blended vegetable oils. *Grain & Oil Science and Technology*, *7*(2), 113–123. https://doi.org/10.1016/j.gaost.2024.05.001

De Luca, M., Ioele, G., Grande, F., Occhiuzzi, M. A., Chieffallo, M., Garofalo, A., & Ragno, G. (2023). Multivariate Curve Resolution Methodology Applied to the ATR-FTIR Data for Adulteration Assessment of Virgin Coconut Oil. *Molecules*, *28*(12). https://doi.org/10.3390/molecules28124661

El-Khalafy, S. H., Hassanein, M. T., Alaskary, M. M., Ramzy, G. H., & Ali, A. I. (2025). Synthesis, characterization, and dielectric properties of bentonite clay modified with (3-chloropropyl)triethoxysilane and Co(ii) porphyrin complex for technological and electronic device applications. *Materials Advances*, *6*(6), 1931–1949. https://doi.org/10.1039/d4ma00982g

Elmosalami, T. A., Kamel, M. M., Tomashchuk, I., Alzaid, M., & Mostafa, M. (2022). *Characterization and Modeling Quality Analysis of Edible Oils Using Electrochemical Impedance Spectroscopy*. *2022*.

Emilia, I., Putri, Y. P., Novianti, D., & Niarti, M. (2021). Pembuatan Virgin Coconut Oil (VCO) dengan Cara Fermentasi di Desa Gunung Megang Kecamatan Gunung Megang Muara Enim. *Sainmatika: Jurnal Ilmiah Matematika Dan Ilmu Pengetahuan Alam*, *18*(1), 88. https://doi.org/10.31851/sainmatika.v17i3.5679

Fang, W., Ren, K., Liu, T., Shang, J., Jia, S., & Jiang, X. (2024). *An evaluation of random forest based input variable selection methods for one month ahead streamflow forecasting*. 1–12. https://doi.org/:10.1038/s41598-024-81502-y

Fatimah, F., & Sangi, M. E. C. (2010). Kualitas Pemurnian Virgin Coconut Oil (VCO) Menggunakan Beberapa Adsorben. *Chemistry Progress*, *3*(2), 65–69.

Gandhi, K., Sharma, R., Seth, R., & Mann, B. (2022). Detection of coconut oil in ghee using ATR-FTIR and chemometrics. *Applied Food Research*, *2*(1), 100035. https://doi.org/10.1016/j.afres.2021.100035

Hamza, M., & Mahmoud, Y. (2026). *Garage-Fabricated, Ultrasensitive Capacitive Humidity Sensor Based on Tissue Paper*.

Havran, P., Cimbala, R., Kurimsk, J., & Kir, J. (2022). *Dielectric Properties of Electrical Insulating Liquids for High Voltage Electric Devices in a Time-Varying Electric Field*. 14–16.

Hwang, J., Choi, K. O., Jeong, S., & Lee, S. (2024). Machine learning identification of edible vegetable oils from fatty acid compositions and hyperspectral images. *Current Research in Food Science*, *8*(December 2023), 100742. https://doi.org/10.1016/j.crfs.2024.100742

Id, J. L. (2024). *Area under the ROC Curve has the most consistent evaluation for binary classification*. https://doi.org/10.1371/journal.pone.0316019

Jermwongruttanachai, P., Pathaveerat, S., & Noypitak, S. (2024). Quantification of the adulteration concentration of palm kernel oil in virgin coconut oil using near-infrared hyperspectral imaging. *Journal of Integrative Agriculture*, *23*(1), 298–309. https://doi.org/10.1016/j.jia.2023.08.002

Luiz, S., Jr, S., Paiter, L., Galvão, J. R., Roque, D. V., & Chaves, E. S. (2015). *Sensor and Methodology for Dielectric Analysis of Vegetal Oils Submitted to Thermal Stress*. 26457–26477. https://doi.org/10.3390/s151026457

Magdas, D. A., Hategan, A. R., David, M., & Berghian-grosan, C. (2025). *The Journey of Artificial Intelligence in Food Authentication : From Label Attribute to Fraud Detection*. 1–31.

Matlala, L., Bokaba, T., Ndayizigamiye, P., Mhlongo, S., Dogo, E., & Dogo, E. (2026). A comparative analysis of ensemble learning models for predicting lapses in investment policies. *Journal of Management Analytics*, *13*(1), 109–138. https://doi.org/10.1080/23270012.2025.2574030

Mb, C., Hebbar, K. B., Ramesh, S. V, Venkatesh, J., Chikkanna, G. S., Prasad, B. N. M., & Harish, B. S. (2022). *Detection of oil adulteration in virgin coconut oil ( VCO ) through physical characterization*. *October*.

Mishra, V., Pratap, S., Singh, M., Singh, V., & Manohar, R. (2024). Heliyon Optical , rheological , and dielectric properties of coconut oil between 100 kHz and 30 MHz. *Heliyon*, *10*(14), e34565. https://doi.org/10.1016/j.heliyon.2024.e34565

Mustapha, A., Ishak, I., Zaki, N. N. M., Ismail-Fitry, M. R., Arshad, S., & Sazili, A. Q. (2024). Application of machine learning approach on halal meat authentication principle, challenges, and prospects: A review. *Heliyon*, *10*(12), e32189. https://doi.org/10.1016/j.heliyon.2024.e32189

Neves, M. D. G., & Poppi, R. J. (2020). Authentication and identification of adulterants in virgin coconut oil using ATR/FTIR in tandem with DD-SIMCA one class modeling. *Talanta*, *219*(May). https://doi.org/10.1016/j.talanta.2020.121338

Nisa, N. H., Masithoh, R. E., Fahri, M., Pahlawan, R., & Wati, A. T. (2025). *Application of machine learning and deep learning to detect adulteration in food flour based on spectroscopy data : a systematic review*. *01005*. https://doi.org/10.1051/bioconf/202519201005

Othman, S., Mavani, N. R., Hussain, M. A., Rahman, N. A., & Mohd Ali, J. (2023). Artificial intelligence-based techniques for adulteration and defect detections in food and agricultural industry: A review. *Journal of Agriculture and Food Research*, *12*(April), 100590. https://doi.org/10.1016/j.jafr.2023.100590

Perera, D. N., Hewavitharana, G. G., & Navaratne, S. B. (2020). *Determination of Physicochemical and Functional Properties of Coconut Oil by Incorporating Bioactive Compounds in Selected Spices*. *2020*.

Piras, C., Hale, O. J., Reynolds, C. K., Jones, A. K., Taylor, N., Morris, M., & Cramer, R. (2021). Speciation and milk adulteration analysis by rapid ambient liquid MALDI mass spectrometry profiling using machine learning. *Scientific Reports*, 1–9. https://doi.org/10.1038/s41598-021-82846-5

Pratiwi, I., & Yunus, M. (2018). *Pemisahan Asam Laurat dari Virgin Coconut Oil ( VCO ) dengan Metode Saponifikasi dan Sonikasi*. *2*(1), 235–239. https://doi.org/10.1155/2020/8853940

Putra, V. G. V., Ngadiono, & Purnomosari, E. (2016). *Pengantar Listrik Magnet dan Terapannya*. http://www.buku-e.lipi.go.id/penulis/drva001/1533693237buku.pdf

Riyanto, S., Sitanggang, I. S., Djatna, T., & Atikah, T. D. (2023). *Comparative Analysis using Various Performance Metrics in Imbalanced Data for Multi-class Text Classification*. *14*(6). https://doi.org/10.14569/ijacsa.2023.01406116

Romero, M., Yuste, S., Ludwig, I., Pedret, A., Jos, M., Maria, R., Salamanca, P., Sol, R., & Rubi, L. (2022). *Phenol metabolic fingerprint and selection of intake biomarkers after acute and sustained consumption of red-fleshed apple versus common apple in humans . The AppleCOR study*. *384*(February). https://doi.org/10.1016/j.foodchem.2022.132612

Roni, K. A., Rifdah, R., Melani, A., Amina Reformis I, A., & Sri, S. M. (2022). Making Virgin Coconut Oil (VCO) With Enzymatic Method Using Pineapple Hump Extract. *International Journal of Science, Technology & Management*, *3*(3), 685–689. https://doi.org/10.46729/ijstm.v3i3.516

Sabnis, S. M., Rander, D. N., Kanse, K. S., Joshi, Y. S., & Kumbharkhane, A. C. (2024). Spectroscopic measurement and dielectric relaxation study of vegetable oils. *Information Processing in Agriculture*, *11*(3), 397–408. https://doi.org/10.1016/j.inpa.2023.04.002

Sairin, M. A., Abd, S., Id, A., Mun, C. Y., Khaled, Y., Zaman, F., & Id, R. (2022). *Analysis and prediction of the major fatty acids in vegetable oils using dielectric spectroscopy at 5 – 30 MHz*. 1–14. https://doi.org/10.1371/journal.pone.0268827

Ströher, D. J., De Oliveira, M. F., Martinez-Oliveira, P., Pilar, B. C., Cattelan, M. D. P., Rodrigues, E., Bertolin, K., Gonçalves, P. B. D., Piccoli, J. D. C. E., & Manfredini, V. (2020). Virgin Coconut Oil Associated with High-Fat Diet Induces Metabolic Dysfunctions, Adipose Inflammation, and Hepatic Lipid Accumulation. *Journal of Medicinal Food*, *23*(7), 689–698. https://doi.org/10.1089/jmf.2019.0172

Sudhakar, A., & Kumar, S. (2023). Understanding the variations in dielectric properties of mustard ( Brassica nigra L .) and argemone ( Argemone mexicana ) oil blends at different temperatures. *Journal of Food Science and Technology*, *60*(2), 643–653. https://doi.org/10.1007/s13197-022-05649-0

Valantina, S. R., Angeline, D. R. P., Uma, S., & Prakash, B. G. J. (2017). PT SC. *Journal of Molecular Liquids*. https://doi.org/10.1016/j.molliq.2017.04.107

Wibisono, A., Sumarno, T., Kunarto, B., & Sani, E. Y. (2021). Pengaruh lama penyeduhan teh hijau (Camellia sinensis L.) berbantu gelombang ultrasonik terhadap aktivitas antioksidan. *Jurnal Mahasiswa Food Tech. Agr. Product, Universitas Semarang*. *Repository Universitas Semarang*, *5*(3), 55–60. http://www.tjyybjb.ac.cn/CN/article/downloadArticleFile.do?attachType=PDF&id=9987

Wijayono, A., & Putra, V. G. V. (2020). Pengukuran Konstanta Dielektrik Udara Pada Perangkat Kapasitor Plat-Sejajar Berbasis Mikrokontroler Arduino Uno. *JIPFRI (Jurnal Inovasi Pendidikan Fisika Dan Riset Ilmiah)*, *4*(1), 13–26. https://doi.org/10.30599/jipfri.v4i1.651