Measurement: Food 22 (2026) 100288

Contents lists available at ScienceDirect

Measurement: Food

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

Classification of cooking oil aroma with high performance using chemometric and E-nose methods

Imam Tazi [a,*], Muthmainnah [a], Wiwis Sasmitaninghidayah [a], Farid Samsu Hananto [a], Arista Romadani [a], Nasirotul Ulya [a], Agung Teguh Wibowo Almais [b], M. Halvi Rahman [c], Andi Setiono [d], Wildan Panji Tresna [d], Yuwana Pradana [d]

[a] Department of Physics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang Jl. Gajayana No. 50 Malang 65144 Indonesia
[b] Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang Jl. Gajayana No. 50 Malang 65144 Indonesia
[c] Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang Jl. Gajayana No. 50 Malang 65144 Indonesia
[d] Research Center for Photonics, National Research and Innovation Agency, KST B.J. Habibie, South Tangerang 15314 Indonesia

ARTICLE INFO

Keywords:
Classification
E-nose
LDA
Backpropagation
Lard

ABSTRACT

Chemometric analysis using Linear Discriminant Analysis (LDA) and Artificial Neural Networks (ANN) was applied to differentiate the aroma profiles of cooking oil samples using an electronic nose. Five types of oils were tested, namely pure lard, pure olive oil, pure palm oil, lard-olive oil mixture, and lard-palm oil mixture. Each oil type was independently tested 40 times, resulting in a total of 200 measurements. The objective was to develop a method to detect lard adulteration in commercial oils. The LDA analysis revealed that the first discriminant function accounted for 59.6% and the second for 34.5% of the total variance. Olive oil was classified with 100% accuracy, palm oil with 90%, while the remaining 10% overlapped with the lard-olive oil mixture (50 : 50 v/v ratio) group. Pure lard and its mixtures were relatively more difficult to classify accurately. Using the backpropagation ANN method, 94% of the test samples were correctly identified. The ANN model employed 50 training data, 100 testing data, and several data points for cross-validation. The results demonstrated stable weight convergence with outputs closely matching the target values, indicating a low level of overfitting. Optimal performance was achieved at the 185th epoch, with a learning rate of 0.001, 30 neurons in the first hidden layer, 10 neurons in the second hidden layer, and a Mean Squared Error (MSE) of 0.0487. Both methods demonstrated high classification accuracy, with ANN achieving 94% and LDA showing variable performance across oil types.

1. Introduction

Cooking oil is an essential daily commodity that functions as a heat transfer medium in food preparation and processing [1]. Based on the raw material, cooking oil consists of two types: vegetable cooking oil, which comes from seed or fruit extracts, and animal cooking oil, which comes from the fats of certain animal parts [2]. Canola, sunflower oil [3, 4], palm oil [5], and olive oil are examples of vegetable cooking oils from industrial products [6–8]. The ingredients for animal cooking oil are beef oil, lard, fish oil, etc [9].

Efforts to mix animal fats in vegetable cooking oil are carried out to add taste and change the color to the cooking oil, which will give rise to new, more affordable prices. Many industries use animal fat from pork because it is inexpensive and abundantly available. Mixing pork fat with vegetable cooking oil raises concerns for some people, especially Indonesians, the majority of whom follow the Muslim religion [10]. In Islam, all parts of the pig are prohibited for consumption. Distinguishing between pure vegetable oils and those adulterated with pork fat is difficult using conventional methods. Therefore, this method can identify the content of cooking oil mixtures simply and quickly. In this study, we

* Corresponding author.
E-mail addresses: tazi@fis.uin-malang.ac.id (I. Tazi), inna@fis.uin-malang.ac.id (Muthmainnah), wiwis_hidayah_87@fis.uin-malang.ac.id (W. Sasmitaninghidayah), faridsamsu@fis.uin-malang.ac.id (F.S. Hananto), arista.romadani@uin-malang.ac.id (A. Romadani), ulyanasirotul@gmail.com (N. Ulya), agung.twa@ti.uin-malang.ac.id (A.T.W. Almais), halvi.rahman@gmail.com (M.H. Rahman), andi033@brin.go.id (A. Setiono), wild004@brin.go.id (W.P. Tresna), yuwa001@brin.go.id (Y. Pradana).

https://doi.org/10.1016/j.meafoo.2026.100288
Received 28 August 2025; Received in revised form 11 March 2026; Accepted 4 April 2026
Available online 5 April 2026

propose the use of an e-nose to identify the content of mixed cooking oils. The electronic nose that is being developed uses 10 types of gas sensors that can smell cooking oil. Data from measurements of the smell of mixed cooking oils need to be processed using mathematical methods. The most common data-processing methods in data classification are the use of chemometrics and artificial neural networks [11].

In the research that has been carried out, the electronic nose output data patterns were processed using chemometric and ANN methods. Both methods are very important for the successful classification of aroma patterns of various mixed cooking oils. Chemometric methods typically employ principal component analysis (PCA) and linear discriminant analysis (LDA), which are classification data processing methods that involve large data matrices [12,13]. This method has proven to be highly effective for solving multivariable data problems. A multivariable reduction process for only two variables can make the classification better and more realistic [14,15].

An artificial neural network (ANN) is a computational method inspired by the biological nerve cells in the human brain [7,16–18]. There are several characteristics of ANNs, including the network architecture (connection patterns between neurons), network training stage (determining connection weights) [19] and activation function [11,20]. An ANN requires initial data as the training memory to recognize a specified pattern [21–23]. Furthermore, several derivative approaches exist within the ANN framework, including the back-propagation algorithm. Backpropagation is a supervised learning method with favorable computational efficiency and is widely used to solve complex problems. This algorithm consists of three main layers: the input layer, hidden layer, and output layer. The study by Darvishi et al. explains that the e-nose technology equipped with eight metal oxide semiconductor (MOS) sensors can detect the adulteration of powdered milk through the analysis of volatile organic compounds (VOC). Utilizing PCA and ANN methods, the research successfully classified pure and adulterated powdered milk samples, particularly in wet samples with significant differences. The ANN model achieved a high accuracy of 95.6 % and effectively detected adulteration. The results confirm that combining e-nose technology with statistical analysis can be a reliable and rapid tool for assessing the quality and purity of powdered milk products [24]. In addition, Wilson et al. applied e-nose technology and employed statistical analysis to detect and identify subterranean termite species based on volatile organic compound (VOC) emissions from termite bodies. The results indicated that each termite species has a unique odor profile. Coptotermes formosanus exhibited significant differences compared to the Reticulitermes species. Through PCA (Principal Component Analysis) and Quality Factor Analysis (QFA), termite species were effectively distinguished, offering a rapid and non-destructive method for monitoring termites in natural environments and buildings [25]. Rasekh et al. demonstrated that the hot air drying method (HAD1a) is the most effective in producing Mentha spicata L. essential oil with the highest yield and best quality, particularly in preserving the contents of carvone, limonene, and carveol. Analysis using GC/MS and Electronic Nose (e-nose), combined with chemometric methods such as PCA, LDA, SVM, and ANN, proved that each drying method provides a distinct aroma profile to mint leaves. The e-nose showed potential as an efficient and rapid tool for essential oil quality monitoring. These findings provide a theoretical basis for selecting the optimal drying method and enhancing the sensory quality of medicinal plant products [26].

Almost all cooking oils have a similar appearance, making them difficult to differentiate visually. Several parameters, such as viscosity, color, and physical and chemical properties, also have significant similarities. The aroma of certain oils (pure or mixed) can be differentiated using an electronic nose with the help of the multivariate ANN method. This study uses high-performance electronic nose detection and performs multivariable data processing by comparing the chemometric-based LDA and ANN-based backpropagation methods. This study makes the following contributions:

1. Know the advantages of the chemometric technique and the ANN backpropagation method in utilizing cooking oil aroma data from the electronic nose.
2. Know the shortcomings of the chemometric technique and the ANN backpropagation method in utilizing cooking oil aroma data from the electronic nose.
3. Know the differences in aroma and types of cooking oil that contain and do not contain pork.

The contents of this article cover the background of using chemometric techniques and artificial neural networks (ANN) to determine the aroma and type of cooking oil, and their application in several studies discussed in the “Related Work” section. The “Method and Data Preparation” section explains the steps involved in chemometric and ANN techniques and data acquisition. The “Results and Discussion” section describes the experimental process of using chemometric techniques and ANN, including the e-nose circuit used to determine the aroma and type of cooking oil and validate the results. Finally, the “Conclusion” summarizes the experimental findings and discusses opportunities for future research.

2. Related work

Research conducted by *Li et al.* differentiated the aroma of different types of Chinese tangerine peel cultivation using a combination of e-nose, e-tongue, and chemometrics [27]. In principle, the e-nose and e-tongue can be used to differentiate certain odors. In addition, the e-nose and e-tongue can also determine the parameters that cause waste pollution in a river, as in the research of *Moufid* et al. which discusses techniques for determining and evaluating the parameters that cause waste pollution in a river [28].

Jiang et al. explained that the results from e-nose can be analyzed using LDA, support vector machine (SVM), KNN, or PCA [29]. Apart from that, the use of e-noses or electronic noses in previous studies still focused on only one object, and no one has discussed comparing two objects. In our research, we analyzed cooking oils including pure oils and lard adulterated mixtures, the results were analyzed using LDA and ANN.

3. Experimental

3.1. Materials and methods

Oil sample material was obtained from supermarkets in Indonesia. Five types of oil samples were analyzed in this study: lard, olive oil, palm oil, lard-olive oil mixture (50:50 v/v ratio), and lard-palm oil mixture (50:50 v/v ratio). For each oil type, 40 independent samples were prepared, resulting in a total of 200 samples. Each sample was then measured once using the electronic nose system to obtain the corresponding sensor response pattern. This number is considered sufficient for data processing in training a chemometric-based classification model. An electronic nose consisting of 10 types of gas sensors (MQ2,

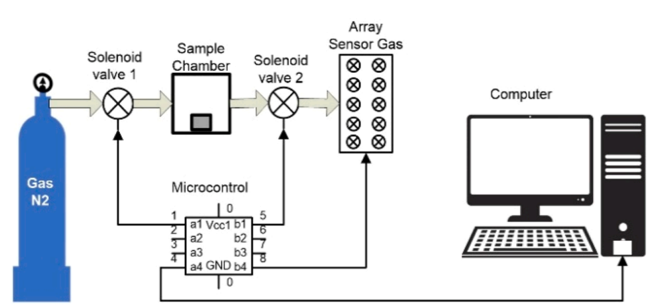

Fig. 1. Experiment design e-nose.

MQ3, MQ4, MQ5, MQ6, MQ7, MQ136, MQ137, MQ138, and MQ303) was used as a data logger to collect the aroma of the cooking oil types (Fig. 1). All Electronic nose data were saved in an Excel format. Classification data processing was performed using LDA and ANN backpropagation method analysis (Fig. 2). Overall, the MQ series of sensors offer a range of capabilities for detecting VOC gases. Each sensor targets specific gases and exhibits varying sensitivities to specific volatile organic compounds (VOCs). Utilizing a combination of these sensors within an electronic nose system can provide more comprehensive and accurate detection of various types of VOCs. Array sensor systems that leverage these non-selective gas sensors enable a wide range of applications in data classification analysis.

3.2. Sample preparation

This study utilized several types of oils as test samples, which were purchased from supermarkets in Indonesia. The samples included lard, olive oil, palm oil, a mixture of lard and olive oil in a 50:50 ratio, and a mixture of lard and palm oil in a 50:50 ratio. Before the aroma measurement, each oil sample was measured to 500 ml and poured into a measuring cup. The oil samples were then placed in an electronic nose (e-nose) chamber and heated to 60 °C to enhance volatile compound release. Once the samples reached the desired temperature, the aroma was measured using the e-nose, equipped with 10 types of gas sensors to detect various volatile components in the oils. For each type of cooking oil, forty independent samples were prepared and measured using the electronic nose system. The measurement process consisted of three main stages: sensing for 30 s, flushing for 30 s, and data sampling at an interval of one data point per second. The data generated from these measurements were then further processed. Data preprocessing involved calculating the area under the curve (AUC) of the sensor response during the sensing phase at steady-state conditions, which was then normalized against the baseline. This step ensured that the generated data were consistent and reliable for further analysis. The measurement data were provided in a 10×200 matrix format in an Excel file.

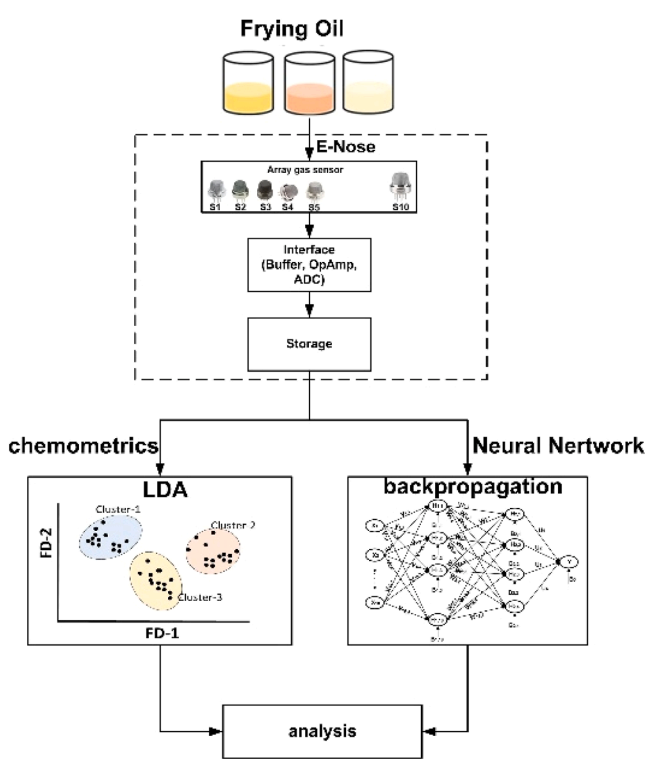

Fig. 2. Frying oil mixture classification research flow chart with an electronic nose.

3.3. Result LDA method

An LDA score plot was used to visualize the aroma patterns of several cooking oils [30–32]. Two hundred measurement data points consisting of five types of cooking oil samples were classified according to their groups using LDA analysis. Statistical analysis was performed using SPSS Statistics. The data matrix had dimensions of 10×200, with a total of 10 variables/sensors (MQ2, MQ3, MQ4, MQ5, MQ6, MQ7, MQ136, MQ137, MQ138, and MQ303). LDA transforms the data matrices into variance and covariance matrices. Next, the eigenvalues from the variance and covariance matrices were obtained as 10 eigenvalues and eigenvectors. The eigenvalues were sorted in descending order, and the corresponding eigenvectors were arranged accordingly. The two largest eigenvalues were used to change the 10×200 matrix to 2×200. The score plot converts the dimensions of the data into two dimensions: discriminant function-1 (FD1) and discriminant function-2 (FD2). These two dimensions explained the data variance of the entire test dataset [33–36]. This test aims to determine the classification pattern of cooking oil based on type and mixture using the parameters in Table I.

3.4. Result artificial neural network method

In this study, the ANN data analysis used a backpropagation algorithm. This algorithm was built using one input layer, two hidden layers, and one output layer. The model results were optimized by creating an architectural model, varying the number of neurons in each layer, varying the learning rate, and observing the MSE. The activation function used was the binary sigmoid function; therefore, input data were normalized to the range [0,1].

$$x' = \frac{0.8(x-b)}{(a-b)} + 0.1 \tag{1}$$

Where
a: data maximum
b: data minimum
x : data
x': data conversion

The development process of the Artificial Neural Network (ANN) consisted of three main stages: training, cross-validation, and testing. The initial step involved defining the input and target data, with the target values presented in Table II. During the training phase, various parameter combinations were evaluated, including different learning rates and numbers of neurons. The best performance was achieved with an MSE value of 0.0487, using 185 epochs, a learning rate of 0.001, 30 neurons in the first hidden layer, and 10 neurons in the second hidden layer. The weights and biases obtained from this optimal training configuration were subsequently applied during the testing phase. The distinction between the training, cross-validation, and testing datasets was used to assess the model's generalization capability and to detect potential overfitting.

Table I
Parameter data.

Cross Validated	Lard	Olive oil	Palm oil	Lard + Olive oil	Lard + Palm oil
Lard	90	0	0	10	0
Olive oil	0	100	0	0	0
Palm oil	0	0	92.5	7.5	0
Lard+ Olive oil	0	0	0	90	10
Lard+ Palm oil	0	0	0	10	90

Table II
Types of frying oils and targets.

Frying Oils	Target
Lard	0
Olive Oil	0.25
Palm Oil	0.5
Lard + Olive Oil	0.75
Lard + Palm Oil	1

4. Discussion

4.1. LDA chemometric

LDA score plots provide an overview of multivariate data classification. Group 1 (lard), Group 2 (olive oil), Group 3 (palm oil), Group 4 (lard : olive oil, 50:50 v/v ratio); and Group 5 (lard and palm oil, 50:50 v/v ratio). FD1 and FD2 accounted for 59.6 % and 34.5 % of the total variance, respectively (Fig. 3). The total value of the first two discriminant functions is 94.1 %. The total value of the last three discriminant functions is 5.9 %. This score plot graph shows that there is 5.9 % missing data information, which is caused by the reduction process from 10 dimensions to two dimensions in LDA. The LDA score plot revealed three distinct data groups. Olive oil and palm oil samples formed distinct, well-separated clusters. Oil groups 1, 4, and 5 (lard + olive oil mixture, and lard + palm oil samples) comprised one group (Fig. 3). Leave-one-out cross-validation was performed to assess classification accuracy for each group. Lard was correctly classified with an accuracy of 90 %, while 10 % of the samples were incorrectly assigned to the lard adulterated olive oil group. The olive oil group could be classified as 100 %. Palm oil was correctly classified with an accuracy of 92.5 %, while 7.5 % of the samples were misclassified as the lard + olive oil group. The lard + olive oil mixture achieved 90 % classification accuracy, with 10 % misclassified as lard + palm oil mixture. Lard + Palm oil was correctly classified with an accuracy of 90 %, while 10 % of the samples were misclassified as the lard + olive oil mixture group (Table I). The aroma profiles of pure olive oil, pure palm oil, and lard exhibit distinct differences (Fig. 3). However, when lard is mixed with either olive oil or palm oil, the aroma profile becomes more similar.

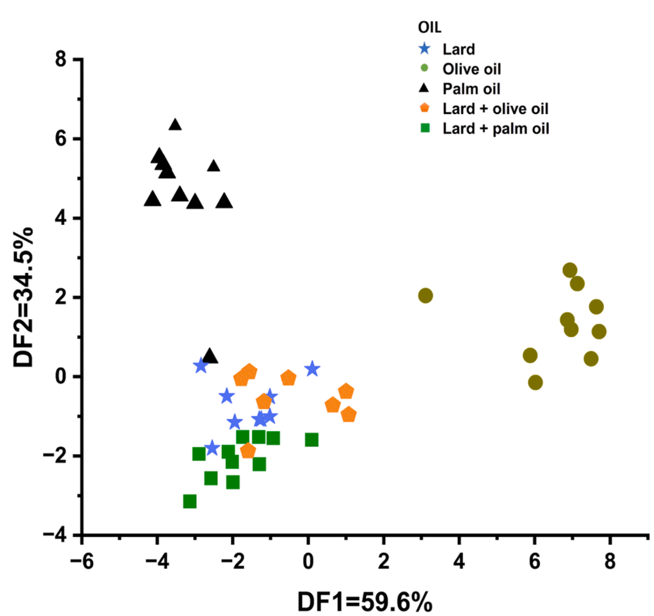

Fig. 3. Score plot of LDA analysis for 5 kinds of Frying oil.

4.2. Pattern recognition using ANN

To reduce the possibility of overfitting, the dataset was randomly divided into training and testing sets during the ANN model development. The first part consisted of 50 data samples used for the training process, while the second part comprised 100 data samples used for testing, along with several additional samples for cross-validation. During the training phase, multiple parameter variations were performed to obtain the weights and biases that yielded the lowest Mean Squared Error (MSE) value. The parameters varied included the number of epochs, learning rate, number of hidden layers, and the number of neurons in each layer. The artificial neural network (ANN) architecture employed consisted of one input layer, two hidden layers, and one output layer. The optimal architecture was represented by the neuron configuration of 10–30–10–1 (Fig. 4). The lowest MSE value obtained was 0.048724, with an accuracy of 94 %, achieved at the 185th epoch and a learning rate of 0.001. The performance graph of the ANN training process under optimal conditions is illustrated by the training data results (Fig. 5).

From (Fig. 5), the calculation operations to determine the output are as follows:

$$H_{1j} = f\left(B_{1j} + \sum_{i=1}^{10} V_{ij}X_i\right) j = (1, 2, \dots 10) \tag{2}$$

$$H_{2j} = f(B_{2j} + \left(\sum_{i=1}^{18} W_{ik}H_{1i})\right) k = (1, 2, \dots 18) \tag{3}$$

$$Y = fB_3 + \left(\sum_{i=1}^{4} U_{im}H_{2i}\right) m = (1, 2, 3, 4) \tag{4}$$

The regression plot obtained at the 185th epoch illustrates the relationship between the actual target values and the network outputs for the training, validation, testing, and overall datasets. The four plots represent: the training data (blue) used to optimize network weights, the validation data (green) for monitoring and preventing overfitting, the test data (red) for assessing model generalization, and the overall data (black) as a composite of all subsets (Fig. 6). The correlation coefficient for the validation data ($R = 0.889$) indicates a strong positive correlation, suggesting that the model performs well in predicting unseen data. The test correlation coefficient ($R = 0.963$) further confirms the model's high predictive accuracy on independent test samples. Visually, the data points in both validation and test plots closely follow the regression line, reflecting the network's reliable alignment between predicted and actual targets. With R values exceeding 0.88 for both validation and testing, the model demonstrates excellent generalization capability and robust performance across non-training data. The Best Validation MSE

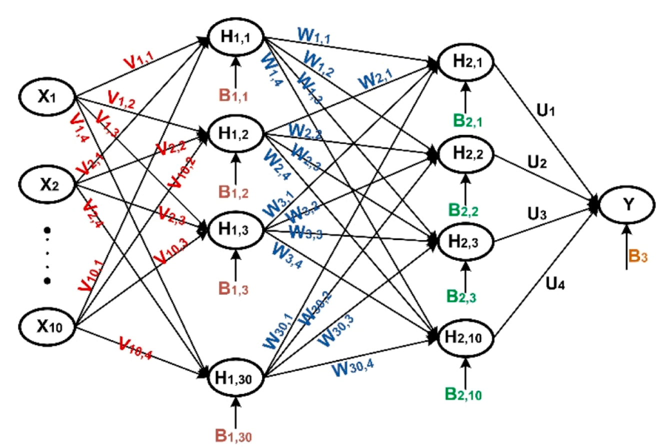

Fig. 4. The results of the ANN architecture using the backpropagation method.

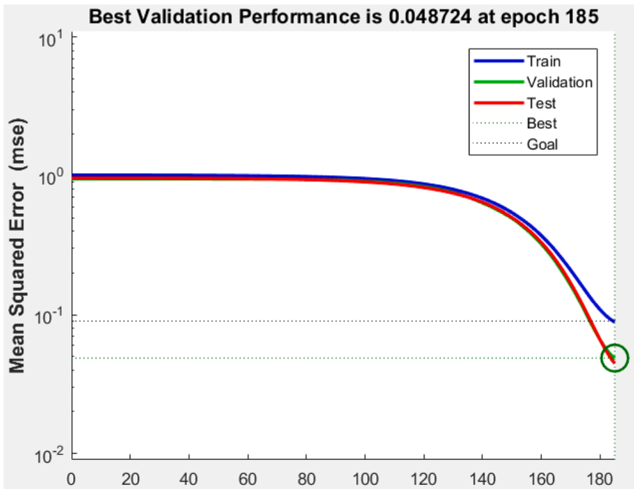

Fig. 5. Graph of best performance results in the training process ANN.

value of 0.0487 is notably low, indicating minimal prediction error. Overall, these results confirm that the ANN model achieves good generalization, high stability between validation and test datasets, and effectively avoids overfitting. Although the training accuracy is moderate, this characteristic reinforces the model's reliability and stability in predicting new data, which is desirable for practical applications.

The network architecture in (Fig. 4) is structured with several transfer functions, because this training function is the fastest and optimal in the training process for each weight (V, W, U) and three biases (B_1, B_2, B_3). The activation function used in all layers is a binary sigmoid function. The training results obtained optimum values with values in the form of weighting and bias matrices, as shown in (Fig. 7–12). $V_{i,j}$ is the weight of the input neuron (Fig. 7), $W_{i,j}$ is the weight of the first hidden layer neuron (Fig. 8), U_i is the weight of the neurons in the output layer (Fig. 9), $B_{1,j}$ is the input layer bias value (Fig. 10), $B_{2,j}$ is the hidden layer value (Fig. 11), and B_3 is the output layer bias (Fig. 12). This matrix was used as a weighting scheme to construct the trained network.

After obtaining the optimal training results, a testing process was performed. The testing process was conducted to determine the extent to which the network could recognize new data patterns. The data used in this test contained 100 data points. From these data, this network was tested and recognized 94 % of the original data.

Majchrzak et al. showed that the integration of an electronic nose with PCA and ANN enables reliable discrimination between adulterated and pure vegetable oils [37]. The difference between this study and the research conducted by Majchrzak et al. lies in the research objective. Majchrzak et al. focused on differentiating adulterated vegetable oils from pure vegetable oils; however, the present study aims to distinguish cooking oils without lard content from those containing lard. For the same tool using an e-nose, the data analysis model uses PCA and ANN to determine the results.

With a low MSE value (0.048724) and high accuracy (94 %), this method demonstrates superior performance compared with the LDA method, as it is highly effective for classification purposes without data reduction. Although the network architecture is relatively complex (10–30–10–1 neurons), potential overfitting was evaluated by comparing the model performance on the training and testing datasets. The results indicate that the testing accuracy (94 %) is comparable to the

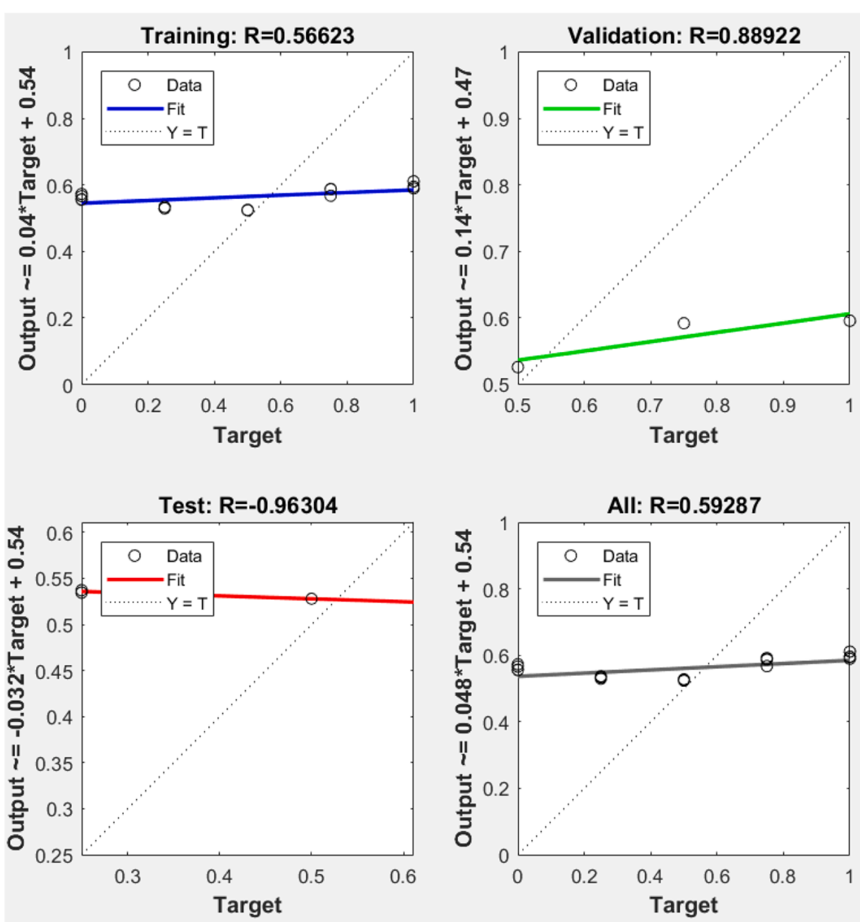

Fig. 6. Artificial neural network training regression plot.

$$\begin{bmatrix}
-0.0297 & -0.0474 & 0.0426 & -0.0138 & 0.0358 & -0.0652 & 0.0235 & 0.0207 & 0.0383 & -0.0323 \\
-0.0149 & 0.0196 & 0.0195 & 0.0284 & 0.0059 & -0.0107 & -0.0092 & 0.01606 & 0.0104 & 0.0079 \\
0.0582 & 0.0351 & 0.0116 & -0.0233 & -0.0179 & 0.0732 & 0.0189 & 0.01678 & 0.0228 & 0.0079 \\
-0.0271 & -0.0393 & 0.0079 & -0.0031 & 0.0070 & -0.0346 & 0.0227 & 0.0069 & 0.0303 & -0.0075 \\
-0.0001 & -0.0236 & 0.0088 & 0.0191 & -0.0202 & 0.0154 & -0.0045 & -0.0317 & 0.0207 & -0.0218 \\
-0.0055 & 0.0239 & 0.0246 & 0.0127 & 0.0073 & 0.0670 & -0.0332 & -0.0050 & -0.0009 & 0.0205 \\
0.0433 & 0.0226 & -0.0295 & -0.0266 & -0.0204 & 0.0299 & 0.0058 & -0.0158 & -0.0306 & 0.0285 \\
-0.0217 & 0.0216 & 0.0071 & -0.0068 & -0.0263 & -0.0078 & -0.0049 & 0.0338 & -0.0116 & 0.0313 \\
-0.0218 & 0.0115 & -0.0116 & -0.0063 & 0.0201 & -0.0039 & 0.0292 & -0.0059 & -0.0146 & 0.0354 \\
-0.0022 & 0.0246 & -0.0038 & -0.0384 & 0.0015 & -0.0173 & 0.0252 & -0.0225 & 0.0037 & 0.0241 \\
0.0012 & -0.0541 & 0.0281 & -0.0029 & 0.0272 & -0.0436 & 0.0176 & -0.0088 & -0.0262 & -0.0020 \\
0.0372 & 0.0271 & 0.0143 & 0.0249 & 0.0226 & 0.0503 & 0.0357 & -0.0013 & -0.0299 & 0.0158 \\
0.0094 & 0.0162 & -0.0245 & 0.0167 & 0.0223 & -0.0294 & 0.0104 & -0.0171 & -0.0226 & -0.0217 \\
0.0307 & 0.0462 & -0.0221 & 0.0226 & 0.0234 & 0.0409 & 0.0201 & -0.0249 & 0.0039 & 0.0352 \\
0.0227 & 0.0310 & 0.0078 & 0.0235 & -0.0101 & -0.0016 & 0.0104 & 0.0119 & -0.0287 & -0.0211 \\
0.0437 & 0.0105 & -0.0068 & -0.0115 & 0.0054 & 0.0574 & 0.0275 & -0.0195 & 0.0083 & -0.0240 \\
0.0050 & 0.0480 & 0.0024 & -0.0064 & -0.0079 & 0.0651 & -0.0200 & 0.0186 & -0.0014 & -0.0206 \\
0.0009 & -0.0322 & -0.0215 & 0.0092 & -0.0226 & -0.0528 & -0.0015 & 0.0337 & -0.0028 & 0.0384
\end{bmatrix}$$

Fig. 7. Input layer weight matrix.

$$\begin{bmatrix}
0.0302 & 0.0109 & 0.0022 & 0.0061 & 0.0170 & 0.0094 & -0.0020 & 0.0193 & 0.0143 & 0.0111 & 0.0088 & 0.0057 & 0.0139 & 0.0259 & 0.0057 & 0.0244 & 0.0210 & 0.0152 \\
0.0765 & 0.1129 & 0.0508 & 0.0627 & -0.0246 & 0.0452 & 0.0090 & 0.0408 & 0.0755 & 0.0708 & 0.0126 & -0.0269 & 0.0446 & 0.0943 & 0.0329 & 0.0871 & 0.1054 & 0.1029 \\
0.0411 & 0.0767 & 0.0908 & 0.0811 & 0.0295 & 0.0844 & 0.0575 & 0.0638 & 0.0714 & 0.0624 & 0.0564 & 0.0294 & 0.0504 & 0.0771 & 0.0561 & 0.0791 & 0.0934 & 0.1129 \\
0.0447 & 0.0967 & 0.0191 & 0.0408 & -0.0090 & 0.0320 & 0.0059 & 0.0260 & 0.0298 & 0.0475 & 0.0146 & -0.0231 & 0.0034 & 0.0568 & 0.0119 & 0.0382 & 0.0707 & 0.0648
\end{bmatrix}$$

Fig. 8. Hidden layer weight matrix.

$$\begin{bmatrix} -0.3943 & -0.4613 & -0.4474 & -0.4258 \end{bmatrix}$$

Fig. 9. Layer output weight matrix.

training accuracy, suggesting that the model does not exhibit significant overfitting and has good generalization capability. The ANN method demonstrated the effectiveness of using an electronic nose and artificial neural networks for classifying cooking oil based on lard content. The classification accuracies reported in this study are based on sample-level classification. Therefore, the dataset consisted of independent observations, reducing the possibility of overfitting due to repeated measurements from the same sample.

Che Man et al. employed a commercial FTIR spectroscopy system combined with chemometric methods to analyze and classify lard and other edible oils based on their molecular composition obtained from infrared spectral data [10]. Although FTIR spectroscopy provides reliable chemical characterization, it primarily reflects bulk molecular information rather than volatile profiles. In contrast, the present study utilizes an electronic nose coupled with LDA and ANN methods, enabling simpler sample handling, faster analysis, and effective classification based on volatile compounds, particularly when dealing with complex oil samples.

5. Conclusions

Classifying the aroma of several types of cooking oils using an electronic nose is an interesting option. The electronic nose can characterize the aroma of several cooking oils quite well, quickly, and easily. This study compares two different data processing techniques in a framework to find the best results for classifying the aroma of different types of cooking oil, especially cooking oil mixed with lard. Based on the LDA method, score plot image classification can distinguish the types of ingredients from palm oil, olive oil, and lard. This score plot still has weaknesses, namely, that it is difficult to classify the types of mixtures of lard adulterated palm oil and lard adulterated olive oil. In the cross-validation data, there are still several types of oil that are not identified according to their groups. The ANN method with the built network architecture can differentiate all cooking oils quite well. The training data process can be obtained quickly with a mean squared error of 0.0487. The resulting weight and bias values can be used to build a smart network architecture that is both stable and reliable. This can be shown by the results of the testing sample data, which had a data match of 94 %. The ANN method has the advantage of being able to draw conclusions about the type of cooking oil directly. Despite the promising results, this study has some limitations. The evaluation was conducted using a limited number of cooking oil types under controlled laboratory conditions, which may not fully reflect real-world variability. Moreover, the relatively complex ANN architecture may introduce a risk of overfitting when applied to larger and more diverse datasets. Future work should focus on expanding the dataset, testing the system under varied environmental conditions, and optimizing the network architecture and validation strategies to enhance model robustness and generalization for practical applications.

$$\begin{bmatrix} -0.0581 \\ 0.1188 \\ 0.1464 \\ 0.2037 \end{bmatrix}$$

Fig. 11. Hidden layer bias matrix.

$$\begin{bmatrix} -1.2164 \end{bmatrix}$$

Fig. 12. Output bias matrix.

[0.0338 0.0739 -0.027 0.0382 0.0593 0.0408 -0.0014 0.0029 0.0043 -0.0077 -0.0048 0.02205 0.0128 0.0417 0.0598 0.0611 -0.0197 0.0324]

Fig. 10. Matrix transpose bias input.

Funding and data availability

Open Access funding enabled and organized by National Research and Innovation Agency (BRIN) (No. 37/II.7/HK/2023). All data will be made available upon request through tazi@fis.uin-malang.ac.id

CRediT authorship contribution statement

Imam Tazi: Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Muthmainnah:** Writing – review & editing, Writing – original draft, Supervision, Project administration. **Wiwis Sasmitaninghidayah:** Writing – review & editing, Validation. **Farid Samsu Hananto:** Writing – review & editing, Writing – original draft, Software, Investigation. **Arista Romadani:** Writing – review & editing, Validation. **Nasirotul Ulya:** Visualization, Validation, Software, Data curation. **Agung Teguh Wibowo Almais:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Investigation, Conceptualization. **M. Halvi Rahman:** Visualization, Validation, Software, Data curation. **Andi Setiono:** Visualization, Validation, Supervision, Funding acquisition. **Wildan Panji Tresna:** Visualization, Validation, Supervision, Software, Investigation, Funding acquisition, Conceptualization. **Yuwana Pradana:** Visualization, Validation, Supervision, Funding acquisition.

Declaration of competing interest

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

Acknowledgment

We extend our deepest gratitude to the Deputy for Research and Innovation Facilitation of the National Research and Innovation Agency (BRIN) Indonesia 2023 for their funding support with research grant contract number 37/II.7/HK/2023, and to the Ministry of Religious Affairs of the Republic of Indonesia for providing facilities for this research. This research could not have been completed without their significant help and support. We also express our thanks to everyone involved for their contributions to the success of this research.

References

[1] I. El Hamzaoui, et al., The design and development of an innovative and environment friendly pilot unit for the transformation of used cooking oils into biodiesel in the City of Ifrane, Morocco, in: 2021 9th International Renewable and Sustainable Energy Conference (IRSEC), IEEE, 2021, pp. 1–8, https://doi.org/10.1109/IRSEC53969.2021.9741209.

[2] J. Milano, et al., Optimization of biodiesel production by microwave irradiation-assisted transesterification for waste cooking oil-calophyllum inophyllum oil via response surface methodology, Energy Convers. Manage 158 (2018) 400–415, https://doi.org/10.1016/j.enconman.2017.12.027.

[3] J. Ulrych, V. Mentlik, Dielectric properties of sunflower, rapeseed and commonly used mineral oil, in: 2016 17th International Scientific Conference on Electric Power Engineering (EPE), IEEE, 2016, pp. 1–4, https://doi.org/10.1109/EPE.2016.7521753.

[4] A. Pavani, A. Hebale, V. Poojary, S. Parulekar, C. Kiran, K. Neeta, Waste sunflower oil as an alternative fuel for diesel engines, in: 2015 International Conference on Nascent Technologies in the Engineering Field (ICNTE), IEEE, 2015, pp. 1–6, https://doi.org/10.1109/ICNTE.2015.7029919.

[5] D. Zhao, et al., Effect of palm oil content and melting point on the freeze-thaw stability of fat substitutes, LWT 188 (2023) 115452, https://doi.org/10.1016/j.lwt.2023.115452.

[6] N. Kishimoto, A. Kashiwagi, Evaluation of filtration on volatile compounds in Virgin olive oils using an electronic nose, in: 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2019, pp. 1–3, https://doi.org/10.1109/ISOEN.2019.8823161.

[7] N. Kishimoto, K. Iwata, A. Utsumi, T. Yajima, Sensory characterization of oils extracted from lacto-fermented olives using an electronic nose, in: 2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2017, pp. 1–2, https://doi.org/10.1109/ISOEN.2017.7968851.

[8] N. Kishimoto, A. Kashiwagi, Prediction of specific odor markers in oil from olive fruit infested with olive scale using an electronic nose, in: 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), IEEE, 2019, pp. 1–3, https://doi.org/10.1109/ISOEN.2019.8823146.

[9] R.L.K. Lokuliyana, H.C. Ambawatte, L.W.T. Rushen, Process optimization for SME level biodiesel production from waste cooking oil, in: 2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS), IEEE, 2016, pp. 1–5, https://doi.org/10.1109/ICIAFS.2016.7946555.

[10] Y.B. Che Man, A. Rohman, T.S.T. Mansor, Differentiation of from other edible fats and oils by means of fourier transform infrared spectroscopy and chemometrics, J. Am. Oil. Chem. Soc. 88 (2) (2011) 187–192, https://doi.org/10.1007/s11746-010-1659-x.

[11] I. Tazi, K. Triyana, D. Siswanta, A.C.A. Veloso, A.M. Peres, L.G. Dias, Dairy products discrimination according to the milk type using an electrochemical multisensor device coupled with chemometric tools, J. Food Meas. Charact. 12 (4) (2018) 2385–2393, https://doi.org/10.1007/s11694-018-9855-8.

[12] I. Tazi, N.L. Isnaini, M. Mutmainnah, A. Ainur, Principal component analysis (PCA) method for classification of beef and pork aroma based on electronic nose, Indones. J. Halal Res. 1 (1) (2019) 5–8, https://doi.org/10.15575/ijhar.v1i1.4155.

[13] Z. Haddi, et al., e-nose and e-tongue combination for improved recognition of fruit juice samples, Food Chem. 150 (2014) 246–253, https://doi.org/10.1016/j.foodchem.2013.10.105.

[14] Q. Ouyang, J. Zhao, Q. Chen, Classification of rice wine according to different marked ages using a portable multi-electrode electronic tongue coupled with multivariate analysis, Food Res. Int. 51 (2) (2013) 633–640, https://doi.org/10.1016/j.foodres.2012.12.032.

[15] B.S. Kang, J.E. Lee, H.J. Park, Electronic tongue-based discrimination of Korean rice wines (makgeolli) including prediction of sensory evaluation and instrumental measurements, Food Chem. 151 (2014) 317–323, https://doi.org/10.1016/j.foodchem.2013.11.084.

[16] N. Nimsuk, Neural network based electronic nose for discrimination of fish sauces, in: 2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), IEEE, 2014, pp. 1–4, https://doi.org/10.1109/ECTICon.2014.6839743.

[17] C. Li, P. Heinemann, R. Sherry, Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection, Sens. Actuat. B Chem. 125 (1) (2007) 301–310, https://doi.org/10.1016/j.snb.2007.02.027.

[18] D. Luo, H.G. Hosseini, J.R. Stewart, Application of ANN with extracted parameters from an electronic nose in cigarette brand identification, Sens. Actuat. B Chem. 99 (2–3) (2004) 253–257, https://doi.org/10.1016/j.snb.2003.11.022.

[19] S.K. Jha, K. Hayashi, R.D.S. Yadava, Neural, fuzzy and neuro-fuzzy approach for concentration estimation of volatile organic compounds by surface acoustic wave sensor array, Measurement 55 (2014) 186–195, https://doi.org/10.1016/j.measurement.2014.05.002.

[20] A. Riul, et al., Wine classification by taste sensors made from ultra-thin films and using neural networks, Sens. Actuat. B Chem. 98 (1) (2004) 77–82, https://doi.org/10.1016/j.snb.2003.09.025.

[21] H.M. Rizal, M. Hariadi, Y.M. Arif, E. Warni, Enhancing flood prediction in urban areas: a machine learning approach for Makassar City, Eng. Technol. Appl. Sci. Res. 15 (2) (2025) 21678–21684, https://doi.org/10.48084/etasr.9402.

[22] R. Haddad, L. Carmel, D. Harel, A feature extraction algorithm for multi-peak signals in electronic noses, Sens. Actuat. B Chem. 120 (2) (Jan. 2007) 467–472, https://doi.org/10.1016/j.snb.2006.02.048.

[23] B. Tudu, A. Jana, A. Metla, D. Ghosh, N. Bhattacharyya, R. Bandyopadhyay, Electronic nose for black tea quality evaluation by an incremental RBF network, Sens. Actuat. B Chem. 138 (1) (2009) 90–95, https://doi.org/10.1016/j.snb.2009.02.025.

[24] P. Darvishi, E. Mirzaee-Ghaleh, Z. Ramedani, H. Karami, A.D. Wilson, Detecting whey adulteration of powdered milk by analysis of volatile emissions using a MOS electronic nose, Int. Dairy. J. 157 (2024) 106012, https://doi.org/10.1016/j.idairyj.2024.106012.

[25] A.D. Wilson, L.B. Forse, Chemical ecology, detection and identification of subterranean termites based on electronic-nose volatile emissions analysis, Environments 11 (1) (2024) 15, https://doi.org/10.3390/environments11010015.

[26] M. Rasekh, H. Karami, M. Kamruzzaman, V. Azizi, M. Gancarz, Impact of different drying approaches on VOCs and chemical composition of Mentha spicata L. essential oil: a combined analysis of GC/MS and e-nose with chemometrics methods, Ind. Crops. Prod. 206 (2023) 117595, https://doi.org/10.1016/j.indcrop.2023.117595.

[27] X. Li, Y. Yang, Y. Zhu, A. Ben, J. Qi, A novel strategy for discriminating different cultivation and screening odor and taste flavor compounds in Xinhui tangerine peel using e-nose, E-tongue, and chemometrics, Food Chem. 384 (2022) 132519, https://doi.org/10.1016/j.foodchem.2022.132519.

[28] M. Moufid, C. Tiebe, N. El Bari, D.A. Hamada Fakra, M. Bartholmai, B. Bouchikhi, Pollution parameters evaluation of wastewater collected at different treatment stages from wastewater treatment plant based on e-nose and E-tongue systems combined with chemometric techniques, Chemom. Intell. Lab. Syst. 227 (2022) 104593, https://doi.org/10.1016/j.chemolab.2022.104593.

[29] H. Jiang, Q. Chen, G. Liu, Monitoring of solid-state fermentation of protein feed by electronic nose and chemometric analysis, Process. Biochem. 49 (4) (2014) 583–588, https://doi.org/10.1016/j.procbio.2014.01.006.

[30] E. Bona, R.S. Dos S. F. da Silva, D. Borsato, D.G. Bassoli, Optimized neural network for instant coffee classification through an electronic nose, Int. J. Food Eng. 7 (6) (2011), https://doi.org/10.2202/1556-3758.2002.
[31] R. Fend, et al., Use of an electronic nose to diagnose mycobacterium bovis infection in badgers and cattle, J. Clin. Microbiol. 43 (4) (2005) 1745–1751, https://doi.org/10.1128/JCM.43.4.1745-1751.2005.
[32] L. Zhang, F.C. Tian, A new kernel discriminant analysis framework for electronic nose recognition, Anal. Chim. Acta 816 (2014) 8–17, https://doi.org/10.1016/j.aca.2014.01.049.
[33] J. Fraden, Handbook of Modern Sensors, Springer International Publishing, Cham, 2016, https://doi.org/10.1007/978-3-319-19303-8.
[34] K. Toko, Biomimetic Sensor Technology, Cambridge University Press, 2000, https://doi.org/10.1017/CBO9780511541179.
[35] H. Lin, et al., Rapid discrimination of Apiaceae plants by electronic nose coupled with multivariate statistical analyses, J. Pharm. Biomed. Anal. 84 (2013) 1–4, https://doi.org/10.1016/j.jpba.2013.05.027.
[36] X. Cetó, M. Gutiérrez-Capitán, D. Calvo, M. del Valle, Beer classification by means of a potentiometric electronic tongue, Food Chem. 141 (3) (2013) 2533–2540, https://doi.org/10.1016/j.foodchem.2013.05.091.
[37] T. Majchrzak, W. Wojnowski, T. Dymerski, J. Gębicki, J. Namieśnik, Electronic noses in classification and quality control of edible oils: a review, Food Chem. 246 (2018) 192–201, https://doi.org/10.1016/j.foodchem.2017.11.013.