RESEARCH ARTICLE

Use of Electronic Nose to Identify Levels of Cooking Cookies

MUHAMMAD RIVAI[1], (Member, IEEE), AND DAVA AULIA[2]
[1]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo-Surabaya, Surabaya 60111, Indonesia
[2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo-Surabaya, Surabaya 60111, Indonesia

Corresponding author: Muhammad Rivai (muhammad_rivai@ee.its.ac.id)

ABSTRACT Currently, the baking of cakes using an electric oven is based on cooking duration. Usually, colors can be used to determine the levels of cooking food. However, many cakes have similar colors at each stage, which cannot be used as indicators of doneness. Through today's technology, the sense of smell can be imitated using a gas sensor combined with artificial intelligence for food quality control. In this study, an electronic nose system was developed to distinguish levels of baking cookies. This process involved 20 gas sensors and 10 classification algorithms based on aroma. The optimization technique based on correlation analysis and distinguishing rate methods was carried out to obtain a small number of sensors that still maintained high accuracy values. Several sensors were eliminated, while the remaining 13 sensors were retained. The selected electronic nose system consisted of 6 gas sensors and convolutional neural networks. It succeeded in distinguishing cooking levels, including undercooked, cooked, and overcooked food, with an accuracy of 90.0%, a precision of 89.7%, a recall of 92.6%, and an F1-measure of 90.2%. This system has the potential to produce a consistent quality of cookies.

INDEX TERMS Aroma, electronic nose, food, levels of cooking cookies.

I. INTRODUCTION

Technological developments in various applications rapidly increased during the industrial revolution. In the cake industry, various roasting devices, such as electric ovens, are designed with an electric power source to build cooking efficiency [1], [2]. Unlike gas ovens, electric ovens have many advantages: a quicker cooking process, stable heat distribution, improved aroma and taste quality [3], [4].

Most electric ovens have a timer system that allows the user to produce food (e.g., cakes, meat, tea, etc.) at the desired cooking level [5]. However, temperature settings and timing systems of electric ovens are still manually adjusted. This will be a problem in determining the baking duration when the cakes have different sizes and ingredients [6]. One must first adjust these to produce cakes with the right cooking level.

Another way to determine baking stages is to analyze color, taste, and texture involving a panelist or sensory test [4], [7]. However, this method will result in a subjective decision. Image processing technology plays a vital role in assessing the baking process's quality characteristics [8], [9], [10]. However, this method will be problematic if the cake has similar coloring throughout multiple cooking levels. Bakery products' quality can be analyzed using spectroscopy [11]. This method has disadvantages, including high operational costs and requiring trained personnel.

In addition to color and taste, baked goods' quality can be determined by the gas or vapor content. Gas chromatography is the standard instrument for quantitatively and qualitatively analyzing volatile organic compounds (VOCs). This method separates compounds in a mixture by injecting a gas sample into the mobile phase, usually called the carrier gas, and passing the gas through the stationary phase [12], [13], [14]. Volatile compounds identified in aldehydes, alcohols, acids, alkanes, and ketones are the main components that contribute

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.

TABLE 1. The ingredients of the cookie.

Ingredients	Values
Butter	100 g
Margarine	150 g
Sugar	50 g
Egg yolk	3 items
Flour	250 g
Cornstarch	25 g
Milk powder	1 tbsp
Chocolate	50 g

to the cookie flavor [15]. Chemical compounds grouped as hydrocarbons are also released while making and storing cookies [16]. In addition to their high cost, these instruments require tens of minutes of analysis time, which is unsuitable for monitoring purposes. Commercial gas sensors may be utilized to enhance the oven's features to overcome this problem [17]. However, this study is only intended to increase safety during cooking.

With today's technology, the human olfactory system can be imitated by a device known as an electronic nose. This system consists of an array of gas sensors and a pattern recognition algorithm. The electronic nose has been widely applied in health [18], [19], food quality control [20], [21], and industrial safety [22] applications. Besides, this method has been applied to tea [23], coffee [24], and bread [25] baking processes. In the latter, the aroma compound content was analyzed during the bread-making process; however, it did not distinguish the cooking levels. The level of gas content released by the bread will change during baking [15], [26]. The electronic nose is unable to identify and quantify specific compounds in the sample. Instead, this method is designed to detect overall patterns rather than individual molecules. This makes it possible to discriminate the cooking levels based on the response pattern of the gas sensor array.

In this study, an electronic nose was implemented to detect levels of cooking cookies during the baking process. The novelty of this research covers determining the appropriate gas sensor array of the electronic nose using the optimization technique based on correlation analysis and distinguishing rate methods. Some artificial intelligence algorithms with various architectures are configured and tested to provide better performance in distinguishing the levels of cooking cookies. The paper is organized as follows. Section I covers this study's background. Section II provides the proposed method and experimental setup. Section III presents the experimental results and discussion. Section IV summarizes the essential results and future works.

II. MATERIALS AND METHODS
A. SAMPLE
The test sample for this study was twenty low-moisture cookies filled with a small piece of chocolate weighing 25 grams each with uniform ingredients. Table 1 shows the cookies' composition. The cooking procedure began with constantly

baking the samples one at a time using an electric oven at 150°C for 45 minutes. Ten experienced assessors and staff of the Nutritional and Food Ingredient section of the Energy and Environmental Laboratory, Institut Teknologi Sepuluh Nopember, served as panelists. Cookies were evaluated for taste, smell, and texture using a five-point hedonic scale, with 1 being very poor and 5 being very good. This lab also measures the moisture content of the cookie samples conducted under Indonesian National Standard SNI 2973:2011. The Mestek WM700A moisture meter is also used for comparison. Color parameters of the upper surface of the cookies were assessed using a WR10QC colorimeter in CIELAB color space. Measurements were carried out on the cookie's surface at least in three different places. The lightness (L*), the green–magenta intensity (a*), and the blue–yellow intensity (b*) values were recorded. The baked samples were considered to produce a distinctive aroma that represented three different cooking levels: undercooked, cooked, and overcooked. As a result, this aroma could be used as an indicator of levels of cooking cookies for the electronic nose. This could be used as a dataset for classification purposes.

B. GAS SENSOR ARRAY
The gas sensor array comprises several semiconductor gas sensors with different characteristics. The main criteria for the gas sensors were high sensitivity to certain gases, stable response, and operation at five volts. The sensing mechanism of semiconductor gas sensor generally can be explained with oxygen adsorption on the surface, as depicted in Figure 1. When the n-type semiconductor material (e.g., SnO_2, TiO_2) is exposed to air, oxygen molecules will be adsorbed on the surface while capturing the free electrons. Specifically, O^- is the most common ion species of adsorbed oxygen at 100-300°C [27]. Then, the electron depletion layer (EDL) width and the potential barrier ($\Delta\emptyset$) height between the grains will be increased, leading to higher resistance of n-type material. The potential barrier is expressed as (1).

$$\Delta\emptyset = \frac{q^2 N_S^2}{2\varepsilon_r \varepsilon_0 n_b} \tag{1}$$

where q is the electron charge, N_s density of surface states, ε_r is the material's relative permittivity, ε_0 is the dielectric constant of vacuum, and n_b is the charge carrier density [28].

When the material is exposed to reducing gas molecules (e.g., ethanol, carbon monoxide), these molecules react with the adsorbed oxygen molecules, which liberate free electrons in the bulk. Then, the EDL width and $\Delta\emptyset$ height decrease again. As a result, this reaction leads to lower resistance of the sensor material.

Table 2 describes detailed information regarding the type of gas sensor used in this study. Most selected sensors were of the metal oxide semiconductor gas sensor type. The gas sensors of the MQ and TGS series are provided by Hanwei Electronics Co., Ltd. and Figaro Engineering Inc., respectively. The gas sensors' responses were represented by conductivity changes when they received gas or vapor samples. Figure 2
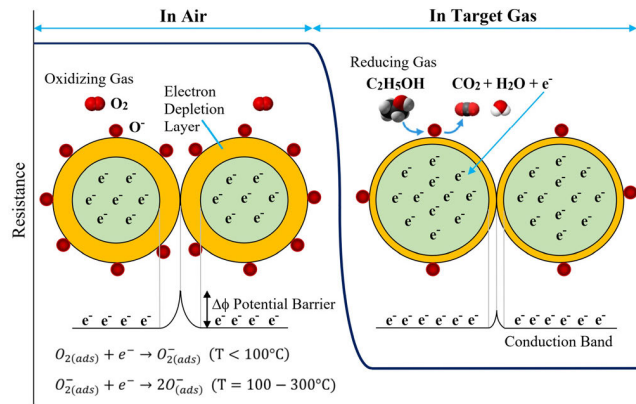
FIGURE 1. The sensing mechanism of n-type semiconductor gas sensor for reducing gas of ethanol.

TABLE 2. Gas sensors and its characteristics.

Sensors	Gases
MQ-2	LPG, propane, methane, alcohol, hydrogen, and smoke
MQ-3	Alcohol and benzene
MQ-4	Methane and LPG
MQ-5	LPG, methane, and hydrogen
MQ-6	Propane, butane, and LPG
MQ-7	Hydrogen, and carbon monoxide
MQ-8	Hydrogen
MQ-9	Carbon monoxide, LPG and methane
MQ-131	Ozone, chlorine, and nitrogen dioxide
MQ-135	Acetone, toluene, and ammonium
MQ-136	Hydrogen sulfide, ammonium, and carbon monoxide
MQ-137	Ammonia, ethanol, and carbon monoxide
TGS813	Hydrogen, isobutane, propane, ethanol, and methane
TGS822	Ethanol, benzene, and hexane
TGS2600	Hydrogen, ethanol, and isobutane
TGS2602	Toluene, hydrogen sulfide, ethanol, and ammonia
TGS2610	Propane, isobutane, methane, hydrogen, and ethanol
TGS2611	Methane, isobutane, hydrogen, and ethanol
TGS2620	Ethanol, hydrogen, isobutane, and carbon monoxide
TGS4161	Carbon dioxide

FIGURE 2. Circuit schematic of the gas sensor [29].

FIGURE 3. Block diagram of electronic nose system for identifying the aroma of cookies.

FIGURE 4. The realization of the electronic nose system for identifying the aroma of cookies.

is a circuit schematic of the gas sensor [29]. This circuit required two voltage sources: heating (V_H) and supply (V_C). Usually, the heater voltage is coupled to the supply voltage for simplicity. V_H can maintain the oxidation–reduction (redox) reactions of the sensor material at a certain temperature. At the same time, V_C will generate a sensor voltage (V_{RS}) whose changes can also be observed in the load resistance (V_{RL}).

C. ELECTRONIC NOSE

The electronic nose system for identifying levels of cooking cookies by aroma was designed by implementing several main components, namely a gas sensor array, an Arduino Uno microcontroller, and a laptop, as depicted in Figure 3. The air filter consisted of silica gel that provided clean and dry air. The electric solenoid acted as a path changer to circulate dry air or bring sample vapor to the 150 mL sensor chamber. The sensor array generated a unique response signal when
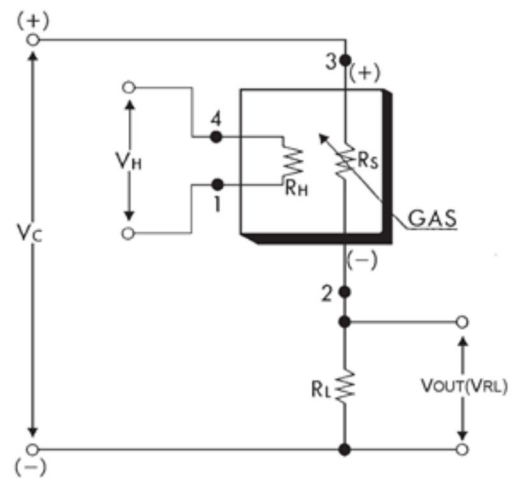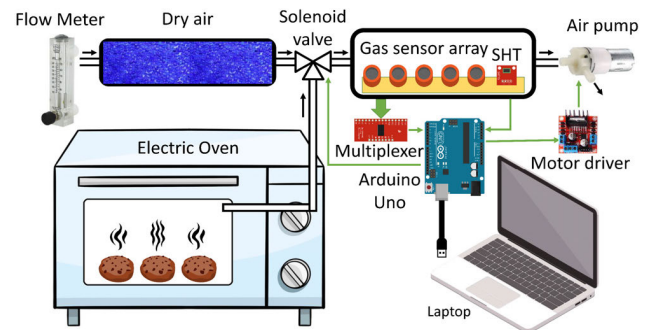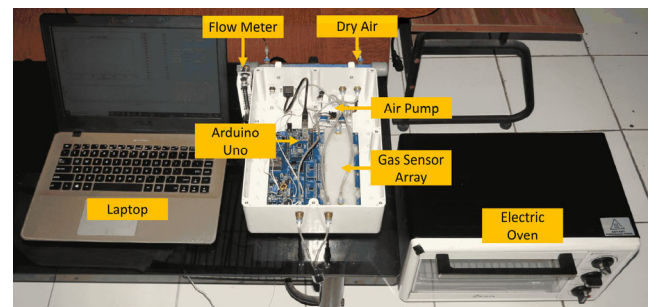
interacting with sample vapors representing each cooking level. The SHT11 module measured humidity and temperature levels in the sensor chamber. The air pump was driven by a DC motor that drew air into the electronic nose system with a flow rate of 100 mL/min. The output voltage from the gas sensor was changed by an analog-to-digital converter (ADC) to be processed on a microcontroller or laptop. The sensor response was sent to the laptop via universal serial bus (USB) communication as a dataset for further processing. Figure 4 is the realization of the electronic nose system for identifying
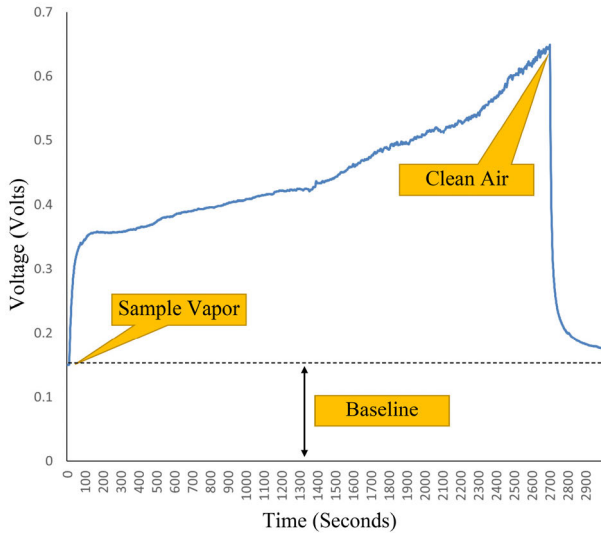
FIGURE 5. The gas sensor response to sample vapor.

FIGURE 6. The concept of compactness and separability in a feature space [32].

cookie aromas. Figure 5 illustrates the gas sensor response to sample vapor in the time domain, where the sensor has a particular baseline value. It can be observed that the sensor response could be used to analyze cooking levels starting from 10 to 2,700 seconds.

D. DATA ANALYSIS

This study involved several types of signals for comparison, and the selected one had the best results. The first type calculates the average value of each sensor data as an unnormalized dataset, which is expressed as (2).

$$\bar{x}_n = \frac{\sum_{T=1}^{k} x_n^T}{k} \tag{2}$$

where k is the number of sample data, n is the sensor number, and x_n^T is the response value of the n-th sensor at time T. The second type is the normalized dataset. This technique may reduce the effects of baseline shifts and intensity variations. This study involves a dataset that is normalized to peak intensity, illustrated in (3).

$$x_{norm} = \frac{x_n}{x_{max}} \tag{3}$$

where x_n is the value of the n-th sensor and x_{max} is the maximum value of the x sensor. The unnormalized and normalized dataset is usually employed to determine the feature that can provide the best performance [30]. Both types were applied to the sensor response, with and without a baseline.

E. OPTIMIZATION OF THE NUMBER OF SENSORS

The selection of the right number of sensors can provide optimal system performance. Determining the selected sensors can involve the role of the coefficient of correlation and cluster analysis. In this study, the sensor array optimization procedure included the following:
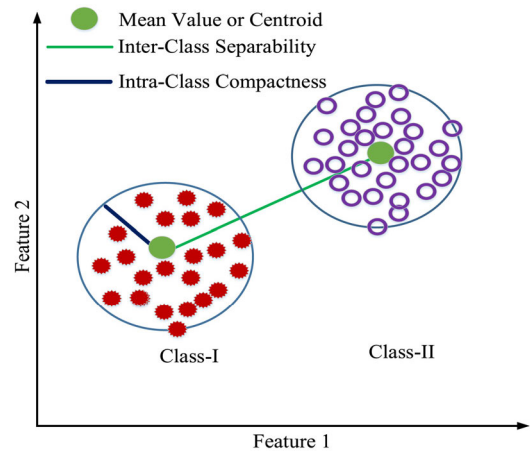 1) Correlation analysis

  2) Distinguishing rate (DR)
  3) Cluster analysis

Correlation analysis was performed by measuring the correlation coefficient between the two sensors to determine the degree of similarity [31]. A high value meant that the two sensors had a strong correlation, indicating that the two responses were similar and thus would be interchangeable. Equation (4) is a calculation of the correlation coefficient.

$$|R_{xy}| = \left| \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2} \right| \tag{4}$$

where x and y represent two different sensors, $\bar{x}$ and $\bar{y}$ are the average values, n is the number of samples, x_i is the i-th data value of the x sensor, y_i is the i-th data value of the y sensor, and $|R_{xy}|$ is the correlation coefficient's absolute value between sensors. For this reason, the sensor's ability to distinguish between classes was tested again by calculating the inter- and intra-class dispersion through the DR method, which was defined as (5).

$$DR = \frac{d(p, q)}{S^2} \tag{5}$$

where $d(p, q)$ is the Euclidean distance between classes p and q, and S^2 is the variance of sensor data from each category. Sensors with smaller DR were removed. Figure 6 is a concept diagram of inter- and intra-class dispersion [32]. Euclidean distance is a calculation to measure the distance between two points, expressed as (6).

$$d(p, q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{6}$$

where n is the number of samples, p and q are the vector values for each class, p_i and q_i are the i-th vector values. The variance was used to determine how far the spread was in the sensor data distribution, formulated as (7).

$$S^2 = \frac{\sum (x_n - \bar{x})^2}{n - 1} \tag{7}$$

Cluster analysis plays an important role in the construction of sensor arrays. In determining the array of sensors, the independence between sensors was considered by measuring the distance between different sensors through cluster analysis. Sensors involved in the same cluster were selected by the sensor with the highest DR. Determination of selected sensors can also be obtained from the principal component analysis (PCA) method extracted from the PCA scores and loadings [33].

F. CLASSIFICATION ALGORITHMS

In this study, several models were applied as pattern recognition algorithms, including support vector machine (SVM), random forest (RF), and linear discriminant analysis (LDA), which were obtained from the scikit-learn library, extreme gradient boosting (XGBoost) were acquired from the XGBoost library, and artificial neural network (ANN), one-dimensional convolutional neural network (1D-CNN), long short-term memory (LSTM), bidirectional long short-term memory (BiLSTM), 1D CNN-LSTM, and 1D CNN-BiLSTM were obtained from the Keras library.

1) SUPPORT VECTOR MACHINE (SVM)

SVM includes supervised learning techniques to solve classification and regression problems in linear and non-linear tasks based on statistical learning theory [34], [35]. This model can separate data into classes by constructing several lines (hyperplanes), which are called SVM decision boundaries. Margins, the distance between hyperplanes, are determined by SVM kernel calculations. The radial basis function (RBF) is a popular kernel based on the Gaussian distribution function [36], which is expressed in (8).

$$K\left(x_i, x_j\right) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \qquad (8)$$

where $\|x_i - x_j\|$ is the Euclidean distance between two points x_i and x_j, and σ^2 is the variance. This study involves several SVM parameters: the RBF kernel, C is 1, and others follow the default. The C function regulates the classification error by adjusting the inclusion of accurate training data in accordance with the margin size of the decision function. When the C value is increased, the decision function aims to create a narrow margin on the decision line in order to distinguish the data classes. This can result in the model overfitting. Nevertheless, when the C value is reduced, it forms a wider margin, making certain data points within specific classes indistinguishable. This phenomenon leads to the model exhibiting underfitting.

2) RANDOM FOREST (RF)

RF is composed of several tree structures, known as ensemble learning, to improve its performance in classification and regression problems [37]. This algorithm is done in parallel, known as the bagging model, where all decision trees give one output, which is then evaluated by a majority voting technique

to produce one final result [38], [39]. The final result of each decision tree is obtained by the split method, which recursively separates the entire population into sub-populations based on their attributes or features [40]. The Gini is the function used to measure a split's quality, which is illustrated in (9).

$$Gini = 1 - \sum_{i=1}^{N} (p_i)^2 \qquad (9)$$

where N is the number of class labels, and p_i is the proportion of i-th class labels. The RF parameters used were the number of trees or n_estimators of 100, the function to measure the quality of a split or criterion was Gini, and the max_depth was set at none. Trees with greater depth can yield a larger volume of data. Similarly, an increased number of trees may yield more favorable outcomes. The remaining parameters were established close to their default values.

3) EXTREME GRADIENT BOOSTING (XGBOOST)

XGBoost has a similar architecture to RF, in which the learning phase of RF is parallel, while XGBoost is sequential [41]. This sequential procedure is called gradient boosting [42], [43], which aims to strengthen weak classifiers' performance. The predicted score of each tree is added up to get the final score. Equation (10) describes the objective function of each round.

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, y_i^{(t-1)} + f_t\left(x_i\right)\right) + \Omega(f_t) \qquad (10)$$

where l is the loss function, $y_i^{(t)}$ is the predicted result of i-th in t-th iteration, and Ω is a regularization term. In this case, the parameters of XGBoost with RF were the same as their respective values. However, XGBoost has an additional parameter, namely the learning rate, which was set at 0.3. Increasing the learning rate might expedite the model's computational speed; nevertheless, this may result in inadequate prediction accuracy as a significant portion of the training data may not be effectively utilized.

4) LINEAR DISCRIMINANT ANALYSIS (LDA)

LDA is a supervised learning algorithm used for classification tasks [44]. It can also be used as a dimensionality reduction technique, which maximizes the separation between classes by finding a set of linear discriminants that maximizes the ratio of between-class to within-class variance [45]. The LDA classifier's performance depends on the sample covariance matrix and class mean vectors, where a larger number of samples can provide sufficient accuracy compared to the number of data dimensions. LDA parameters include the solver type using singular value decomposition, number of components to 'none', and others being defaulted. The solver type does not calculate the covariance matrix. As a result, this solver is suggested for datasets that possess a substantial number of features.

5) ARTIFICIAL NEURAL NETWORK (ANN)

ANN is a computing network system inspired by biological neural networks. This model comprises three layers, including input, hidden, and output, each with its function [46]. ANN has two different methods, namely feed-forward and back-propagation [20]. The former is a one-way computation that is applied in the testing phase. In contrast, the latter can fix the weight by recursively returning its value to the previous layer, which is used in the training phase [22]. In this study, the classification of three cooking levels involved three hidden layers, each consisting of 128 neurons with the rectified linear unit (ReLU) activation function. ReLU has a gradient of 1 for positive and 0 for negative values. This characteristic renders ReLU a more favorable approach for mitigating the issue of vanishing gradients. In contrast, the neurons in the output layer utilize the Sigmoid activation function for binary-class classification and the Softmax activation function for multi-class classification. Utilizing multiple layers in a neural network is advantageous for generalization because it can learn the intermediate features between the input data and the high-level categorization. Selecting fewer neurons can result in underfitting and a substantial statistical bias. On the other hand, choosing an excessive number of neurons can result in overfitting, increased variance, and prolonged network training time. During the training phase, the batch size was configured to 25, and the number of epochs was set at 5000.

6) ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK (1D-CNN)

CNN has a convolution layer consisting of filters or kernels that extract specific features from the input data. CNN can be applied as a classification in the field of signal processing, which is termed 1D-CNN [18], [19]. The proposed CNN architecture consists of three convolutional layers with 16, 64, and 32 filters, respectively, with kernel sizes set to 2, 1, and 1, respectively, and a ReLU activation function combined with two neural hidden layer networks of 128 neurons, each fully connected. Expanding the filter size of a CNN has the potential to enhance its accuracy through several means. The size of the kernel is a critical factor in influencing the performance of CNN. A reduced kernel size will enable the capturing of more intricate elements within the data. Conversely, employing a bigger kernel size will enable the detection and extraction of larger, more intricate features.

7) LONG SHORT-TERM MEMORY (LSTM)

LSTM is an algorithm that has a good ability to predict information in the time domain. This algorithm is a development of a recurrent neural network (RNN) that can recall a collection of information stored for a long time while deleting no longer relevant data [47]. This experiment employed an LSTM architecture consisting of two LSTM layers, called stacked LSTM, of 128 memory cells integrated with the ReLU activation function. The stacked LSTM model is a

variant incorporating multiple LSTM layers, each consisting of multiple memory cells. The utilization of stacked LSTM hidden layers enhances the depth of the model, hence justifying its classification as a deep learning technique with improved accuracy.

8) BIDIRECTIONAL LONG SHORT-TERM MEMORY (BILSTM)

BiLSTM, a variant of LSTM, can capture information from two directions to produce more meaningful output that can increase the classification level [48]. The BiLSTM architecture was implemented with one bidirectional LSTM layer of 128 memory cells.

9) 1D CNN-LSTM AND 1D CNN-BILSTM

The combination of CNN and LSTM or BiLSTM can improve performance in making predictions and processing high-dimensional data [49]. This is supported by a convolution layer as feature extraction and a hidden layer that can remember information data and optimize valuable data [44]. These two models combine the architecture of CNN with LSTM and CNN with BiLSTM.

G. PERFORMANCE EVALUATION

The electronic nose system's performance needed to be measured to find out how well it could classify aroma data. The confusion matrix, one of the statistical measurement methods, was applied in this study. This method evaluates the system by providing performance variables, including accuracy, precision, recall, and F1-measure, based on several indicators, such as true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [50]. The performance variables were as follows:

1) Accuracy (11) measured the ratio of correct recognition results to the entire dataset.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (11)$$

2) Precision (12) estimated how many of the positive predictions were correct.

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

3) Recall can also be called Sensitivity (13) calculated how many positive cases the classifier correctly predicted over all the positive instances in the data.

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

4) F1-measure (14) assessed the predictive skill of a model by elaborating on its class-wise performance.

$$F1 - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (14)$$
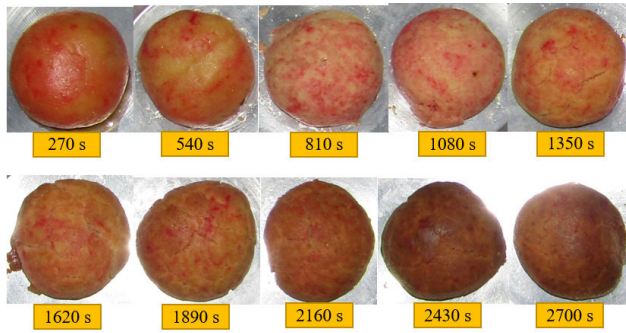
FIGURE 7. The appearance of cookies with different roasting time.

TABLE 3. Characteristics for each roasting time of cookies.

Roasting time (seconds)	Color			Moisture (%)	Taste	Smell	Texture
	L*	a*	b*				
270	58.9	11.1	42.4	12.8	2.4	2.8	2.5
540	58.5	10.9	43.8	12.2	2.0	2.6	2.5
810	61.2	14.0	47.7	8.8	2.3	2.7	2.1
1,080	57.6	18.4	49.2	4.4	3.5	3.1	2.7
1,350	51.3	12.4	37.8	3.4	2.6	2.9	3.1
1,620	51.4	19.5	47.4	3.2	2.7	3.0	3.0
1,890	48.9	21.2	54.1	2.6	2.6	2.9	3.1
2,160	42.8	20.5	49.4	2.2	2.1	3.1	3.0
2,430	39.3	19.8	47.8	2.0	1.6	2.1	2.7
2,700	34.6	19.4	42.7	1.2	1.5	2.0	2.7

TABLE 4. The levels of cooking cookies.

Levels	Time (seconds)
Undercooked	0 – 540
Cooked	1,620 – 1,890
Overcooked	2,430 – 2,700

III. RESULTS
A. LEVELS OF COOKING COOKIES
This study established a procedure for determining proper cooking levels, in which as many as ten samples were baked simultaneously and taken out of the oven one at a time in 270-second increments. Cookies produced in this experiment were then tested using a colorimeter, moisture meter, and panelists. Figure 7 shows the appearance of cookies with different roasting time, while the properties of each cooking level are shown in Table 3. Color is the quality attribute most consumers consider in product acceptance. Initially, roasting cookies had an L* value of 58.9, which tends to be brighter. This intensity decreases over the roasting time with a darker appearance. The positive value of the average of a* 16.7 and b* 46.2 characterizes cookies as red and yellow, respectively. Cookie moisture content after cooking for more than 1,080 seconds is below 7%. The low water content value is an attribute considered to play a major role in the crispness [51]. The panelists have a high organoleptic assessment of taste, smell, and texture when the roasting time ranges from 1,350 to 2,160 seconds. Overall, this shows undercooked cookies were baked for less than 540 seconds. Meanwhile, the cooked
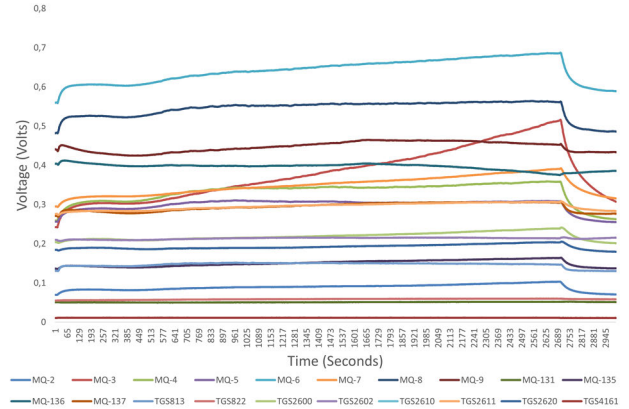

FIGURE 8. The response of the sensor array to the sample vapor.

TABLE 5. Comparison of system performance to each type of sensor response.

Type of Sensor Response	LDA F1-Measure (%)
Unnormalized	96.4
Normalized	97.0
Unnormalized without a baseline	97.6
Normalized without a baseline	93.5

cookies were baked for 1,620 to 1,890 seconds, and the overcooked cookies for more than 2,430 seconds, as shown in Table 4.

B. CLASSIFICATION OF LEVELS OF COOKING COOKIES
Every measurement was performed by placing a cookie in the oven. Each cookie produced a specific aroma at varying points in the baking process. The electronic nose drew sample vapor from the oven through a small hose. The vapor was delivered into the chamber and measured by an array of gas sensors to obtain an aroma pattern. Figure 8 shows the sensor array's response, consisting of 20 gas sensors to the sample vapor. Each sensor had a response that increased with the cooking levels. This indicated that the higher the temperature, the higher the vapor concentration emitted by the sample. In the experiment, twenty cookies were baked to produce a dataset for classification purposes.

The next step was to test several types of responses: unnormalized, normalized, as well as with and without a baseline. Table 5 compares the system performance results for each type of sensor response using the LDA identification algorithm. As a result, unnormalized responses without a baseline had the highest F1-measure value of 97.6%, which was used as a dataset for further processing. Figure 9 shows the sensor array's response pattern for unnormalized responses without a baseline for each cooking level. Of the three-sensor array response patterns, several sensors had similar values for each cooking level. Figure 10 illustrates the classification results using LDA. Each cooking level could be appropriately separated, indicating that the gas sensor array can provide a unique sensor response pattern for each level.
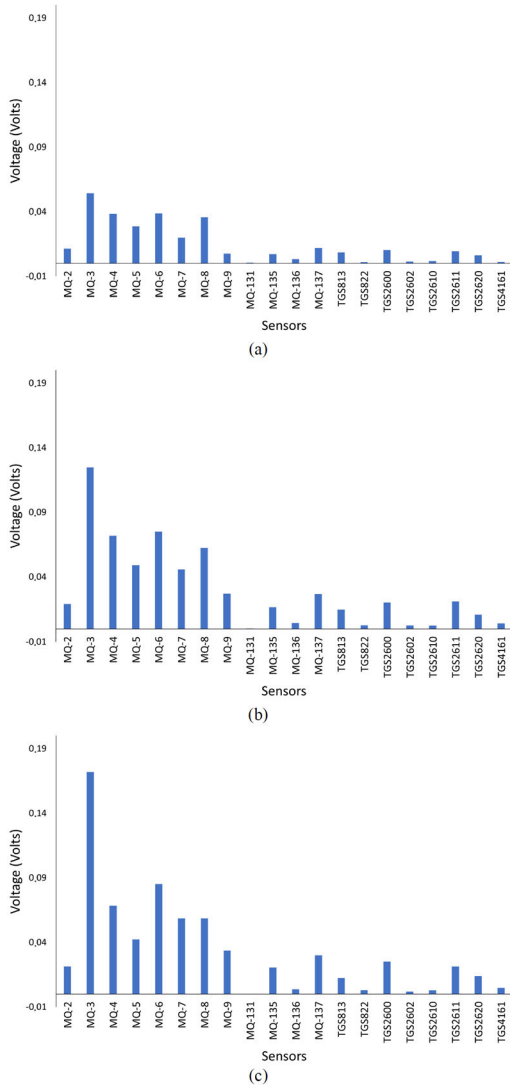
FIGURE 9. The response pattern of the sensor array for unnormalized without a baseline for each cooking level: (a) undercooked, (b) cooked, and (c) overcooked.

FIGURE 10. Visualization of classification results using LDA.

TABLE 6. The sensor pair correlation coefficients for each cooking level.

Undercooked		Cooked		Overcooked	
Sensors	|Rxy|	Sensors	|Rxy|	Sensors	|Rxy|
MQ-2 - MQ-3	0.884	MQ-4 - MQ-5	0.939	MQ-2 - MQ-6	0.872
MQ-2 - MQ-6	0.893	MQ-4 - MQ-8	0.954	MQ-4 - MQ-5	0.941
MQ-2 - MQ-7	0.899	MQ-5 - MQ-8	0.917	MQ-4 - MQ-8	0.965
MQ-3 - MQ-7	0.906	MQ-9 - MQ-137	0.853	MQ-5 - MQ-8	0.920
MQ-4 - MQ-5	0.943	TGS2600 - TGS2620	0.982	TGS2600 - TGS2620	0.969
MQ-4 - MQ-8	0.901				
MQ-5 - MQ-6	0.861				
TGS2600 - TGS2620	0.969				

C. SENSOR ARRAY OPTIMIZATION BASED ON CORRELATION COEFFICIENT AND DR

The electronic nose, consisting of twenty gas sensors, was able to distinguish each cooking level properly. However, it was necessary to optimize the sensor array to obtain a more optimal number of sensors, as well as make it a compact size and at a low cost. The first step was to perform calculations using the correlation analysis method for each sensor over the others. Table 6 shows the results of the sensor pair correlation coefficients for each cooking level. The correlation coefficients above 0.85 were considered candidate sensors selected for elimination. Most highly correlated sensor pairs were undercooked due to the low vapor concentration. The DR for each sensor is shown in Table 7 and ranked for the evaluation stage. The TGS2610 sensor had the first rank with a DR of 175,034.14, indicating the sensor had a good ability
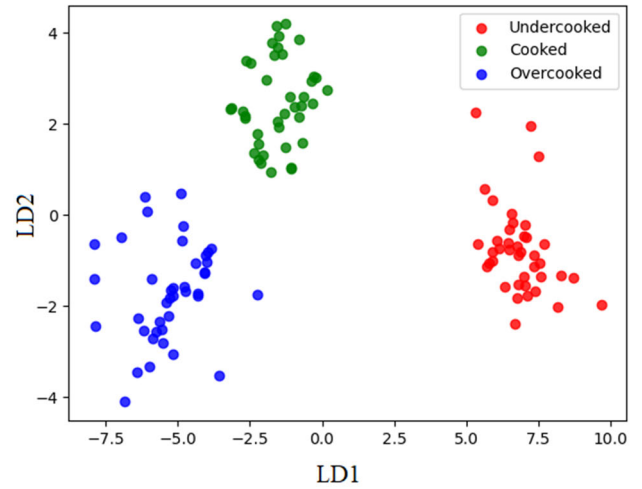
to distinguish between the cooking levels. Besides being able to respond to the volatile compounds produced by cookies during roasting, this sensor has the smallest data distribution value for each cooking level, showing its highest precision level. Meanwhile, the MQ-9 sensor had the lowest rank with a DR of 456.29, exhibiting poor capabilities.

The next step was to evaluate the sensor pairs based on the DR and correlation coefficient. For example, the MQ-2 and MQ-6 sensors for the overcooked level, shown in Table 6, had DRs of 5,646.21 and 1,837.88, respectively, as shown in Table 7. For this reason, the MQ-6 sensor had to be removed because it had a smaller DR, while the MQ-2 sensor was retained to be re-evaluated with other sensors. The final sensor selection process results are shown in Table 8. Several sensors were eliminated, while the remaining 13 sensors were retained, including MQ-2, MQ-5, MQ-131, MQ-135, MQ-136, MQ-137, TGS813, TGS822, TGS2602, TGS2610, TGS2611, TGS2620, and TGS4161. Apart from having a significant contribution in differentiating cooking levels, these selected sensors have responses to chemical compounds released by cookie samples with different sensitivities and limits of detection according to their data sheets.

TABLE 7. The rank of the DR of each sensor for the cooking levels.

Sensors	DR Values
TGS2610	175,034.14
TGS822	33,364.11
MQ-131	22,444.71
MQ-2	5,646.21
TGS2602	2,717.13
MQ-7	2,680.88
TGS813	2,503.41
MQ-137	2,425.83
MQ-6	1,837.88
MQ-135	1,751.79
TGS4161	1,743.15
MQ-5	1,521.37
TGS2620	1,069.84
MQ-4	955.88
TGS2600	900.61
TGS2611	834.50
MQ-8	815.79
MQ-3	649.28
MQ-136	509.56
MQ-9	456.29

TABLE 8. Gas sensor optimization based on correlation coefficient and DR.

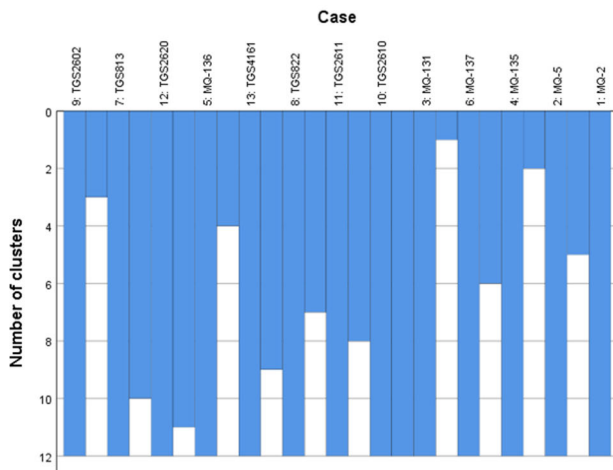| Cooking Levels	Retained Sensors	Eliminated Sensors
Undercooked	MQ-2, TGS2620	MQ-3, MQ-4, MQ-5, MQ-6, MQ-7, MQ-8, TGS2600
Cooked	MQ-5, MQ-137, TGS2620	MQ-4, MQ-8, MQ-9, TGS2600
Overcooked	MQ-2, MQ-5, TGS2620	MQ-4, MQ-6, MQ-8, TGS2600

FIGURE 11. Icicle plot of the averaged sensor response groupings.

D. SENSOR ARRAY OPTIMIZATION BASED ON CLUSTER ANALYSIS AND DR

The remaining sensors were grouped as the selected sensors by involving cluster analysis and DR. For this purpose, an icicle plot diagram was needed, performed in SPSS 25.0 software, shown in Figure 11. Interconnected sensor histograms were aggregated into one group, while the

TABLE 9. Gas sensor optimization based on cluster analysis and DR.

Cluster Numbers	Selected Sensors
2	TGS2610, MQ-2
3	TGS2610, MQ-137, MQ-2
4	TGS2602, TGS2610, MQ-137, MQ-2
5	TGS2602, TGS813, TGS2610, MQ-137, MQ-2
6	TGS2602, TGS813, TGS2610, MQ-137, MQ-5, MQ-2
7	TGS2602, TGS813, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
8	TGS2602, TGS813, TGS822, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
9	TGS2602, TGS813, TGS822, TGS2611, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
10	TGS2602, TGS813, TGS4161, TGS822, TGS2611, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
11	TGS2602, TGS813, MQ-136, TGS4161, TGS822, TGS2611, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
12	TGS2602, TGS813, TGS2620, MQ-136, TGS4161, TGS822, TGS2611, TGS2610, MQ-137, MQ-135, MQ-5, MQ-2
13	TGS2602, TGS813, TGS2620, MQ-136, TGS4161, TGS822, TGS2611, TGS2610, MQ-131, MQ-137, MQ-135, MQ-5, MQ-2

disconnection ones were collected into another. For instance, to produce six sensors, a straight line was made through six clusters on the histogram to obtain the sensor candidates, which included TGS2602, (TGS813/TGS2620/MQ-136), (TGS4161/TGS822/TGS2611/ TGS2610/MQ-131), (MQ-137/MQ-135), MQ-5, and MQ-2. The highest DR sensor was chosen to represent the sensors in its cluster. In this case, the DR values for each sensor were: (TGS2602, 2,717.13), ((TGS813, 2,503.41)/(TGS2620, 1,069.84)/(MQ-136, 509.56)), ((TGS4161, 1,743.15)/(TGS822, 33,364.11)/ (TGS2611, 834.50)/(TGS2610, 175,034.14)/(MQ-131, 22,444.71)), ((MQ-137, 2,680.88)/(MQ-135, 1,751.79)), (MQ-5, 1,521.37), and (MQ-2, 5,646.21). As a result, the TGS813, TGS2610, and MQ-137 sensors were selected and combined with the remaining single sensors, namely TGS2602, TGS813, TGS2610, MQ-137, MQ-5, and MQ-2. Table 9 shows the complete results of the sensor array optimization based on the cluster analysis and DR. All sensor clusters will be tested to obtain the least number of sensors that still have high performance.

E. DETERMINATION OF THE BEST CLASSIFICATION ALGORITHM

All sensor array clusters in Table 9 were tested with ten classification algorithms. Table 10 describes the performance of 10 classifiers for each sensor cluster for three cooking levels. Assuming the best classifier is F1-measure with a baseline above 90%, LDA, XGBoost, ANN, and CNN are the best candidates. Therefore, CNN was chosen as the best algorithm because it provides the fewest number of sensors, namely six. This was necessary for system simplicity while still maintaining high performance.

For the undercooked and cooked levels, all classifiers met the performance criteria shown in Table 11. All classifiers

TABLE 10. F1-measure value of classifiers for each number of sensors for three cooking levels.

Number of Sensors	LDA	SVM	RF	XGBoost	ANN	CNN	LSTM	BiLSTM	CNN-LSTM	CNN-BiLSTM
2	75.2	71.4	69.8	70.2	75.0	72.8	74.6	74.6	71.6	71.2
3	72.4	70.6	71.6	70.2	67.8	68.3	62.8	68.2	65.2	67.0
4	70.8	66.4	71.0	66.2	69.0	74.2	65.0	73.0	70.4	66.6
5	80.0	74.2	80.0	76.6	76.5	81.2	76.6	70.4	71.0	73.0
6	90.6	78.0	86.0	87.2	88.3	94.3	63.2	59.6	69.0	75.6
7	92.2	84.2	84.2	87.8	92.8	94.3	67.8	65.4	75.6	77.2
8	96.0	83.8	88.8	91.6	94.2	94.3	67.2	66.8	83.4	79.4
9	95.2	80.0	89.6	89.6	94.3	94.7	72.6	64.8	80.4	78.6
10	94.8	77.4	88.4	90.2	93.8	94.7	65.4	60.6	77.0	76.8
11	96.0	77.6	89.8	91.8	94.7	94.8	73.6	70.0	76.8	79.2
12	95.6	83.2	85.8	89.6	95.0	95.0	66.2	57.0	79.8	68.8
13	94.2	83.2	88.8	90.4	96.0	95.7	68.6	59.6	76.8	68.8

TABLE 11. F1-measure value of classifiers for each number of sensors for two cooking levels.

Number of Sensors	LDA	SVM	RF	XGBoost	ANN	CNN	LSTM	BiLSTM	CNN-LSTM	CNN-BiLSTM
2	91.7	94.5	94.0	94.0	97.0	97.0	97.0	96.25	97.0	94.75
3	91.8	94.0	94.5	94.0	94.0	94.0	96.25	94.0	94.75	93.25
4	94.5	96.5	95.5	95.0	95.5	97.0	97.75	96.25	97.0	94.75
5	97.6	98.2	95.8	94.6	94.75	94.75	97.0	94.75	93.25	94.75
6	97.0	96.4	95.8	95.2	94.0	94.75	97.0	95.5	94.0	94.0
7	97.0	94.6	98.2	95.0	98.5	98.5	97.0	96.25	98.5	96.25
8	97.6	94.6	100.0	99.4	97.0	96.25	95.5	97.0	98.5	94.75
9	100.0	95.8	100.0	99.4	96.75	96.25	93.0	95.5	97.0	96.25
10	100.0	97.6	98.8	100	98.5	98.5	98.5	97.75	98.5	98.5
11	100.0	96.4	100.0	100	98.5	98.5	97.0	96.25	97.75	96.25
12	97.6	95.2	100.0	100	94.75	95.5	93.25	93.25	95.5	94.75
13	100.0	96.4	100.0	100	100.0	100.0	100.0	100.0	100.0	100.0

TABLE 12. Characteristics for cookies produced by an electric oven equipped with the electronic nose.

Temperature setpoints (°C)	Color L* a* b*	Moisture (%)	Taste	Smell	Texture	Roasting time (seconds)
150	51.4 20.0 49.6	3.4	2.9	3.2	3.5	1,777
175	49.8 18.7 48.0	1.4	3.2	3.5	3.4	1,569
200	51.9 19.6 51.3	3.0	3.3	3.4	3.6	1,482

had a high performance for each number of sensors, with an average F1-measure above 91%. Regarding the smallest number of sensors (i.e., two sensors), ANN, CNN, LSTM, and CNN-LSTM had the same highest value. The simplicity of the ANN architecture was preferable for implementation into a low-level microprocessor system.

F. EVALUATION OF SELECTED ELECTRONIC NOSE SYSTEMS

In this study, the selected electronic nose system involved six gas sensors, consisting of TGS2602, TGS813, TGS2610, MQ-137, MQ-5, MQ-2, and CNN classifiers. This optimum sensor array can provide a specific response pattern to the volatile compounds released by cookies, especially alcohols and hydrocarbons, for each baking level. The CNN algorithm can still provide a discrimination level of 94.3%, according to Table 10. In the next experiment, this electronic nose was then tested for twenty new samples with random and unknown

cooking levels at an oven temperature of 150°C. In this case, the cooking levels were confirmed by the majority decision of the panelists after these measurements and predictions. Figure 12 shows that this system can predict the cooking levels with an accuracy of 90.0%, a precision of 89.7%, a recall of 92.6%, and an F1-measure of 90.2%.

The other test was needed to determine the system's performance in predicting levels of cooking cookies at varying temperatures in an electric oven to show its robustness. Fifteen samples with the material composition shown in Table 1 are 25 grams with a uniform size. Every five samples were tested at different temperature setpoints, namely 150°C, 175°C, and 200°C. The electronic nose system gave off an alarm when the cookies reached the cooked level. The cookies' attributes produced by the electric oven equipped with the electronic nose are shown in Table 12. It can be observed that the three samples had similar color characteristics when cooked; they had the average values of L*, a*, and b*, which are 51.0±1.5,
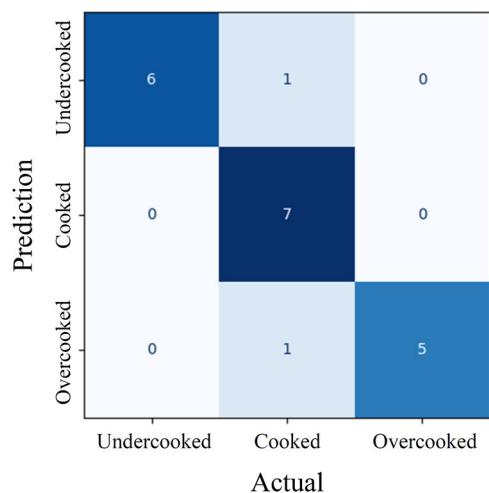
FIGURE 12. The electronic nose prediction against levels of cooking cookies.
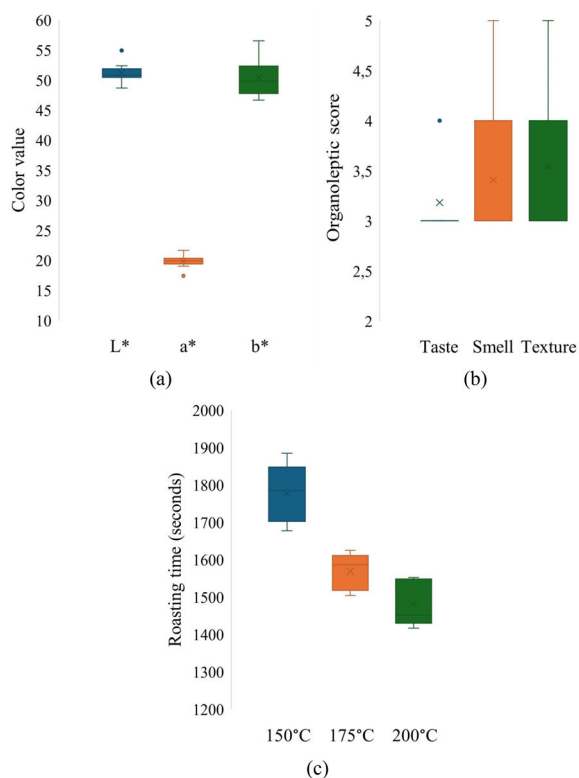

FIGURE 13. Box plots associated with the cookie quality assessment under different oven temperature setpoints: (a) color value, (b) organoleptic score, and (c) roasting time.

19.4±1.9, and 49.6±2.1, respectively. All moisture contents are below 7 %, indicating the appropriate cookies' crispness. Organoleptic evaluations, including taste, smell, and texture, have an average hedonic score of more than 3.3, which expresses a fairly good acceptance. The samples baked at a temperature setpoint of 150°C took an average of about 1,777 seconds to cook. The samples baked at 175°C required

a shorter average cooking time, around 1,569 seconds. Meanwhile, the sample baked at a temperature setpoint of 200°C required the shortest average cooking time at about 1,482 seconds. The higher the oven temperature, the faster the electronic nose gave an alarm indicating that the sample had reached the cooked level.

Figure 13 shows box and whisker plots indicating the robustness of cookie quality assessment under different oven temperature setpoints. Both color values and sensory tests have a minimum number of outliers. Each box in the roasting time has a different mean value and almost does not overlap. This proved that the electronic nose system could identify aroma and predict the levels of cooking cookies, even at different temperatures.

IV. CONCLUSION
In this study, an electronic nose system was developed to classify the cooking level of cookies in an electric oven. Optimizing the number of gas sensors involved correlation analysis and distinguishing rate methods. Several artificial intelligence algorithms were used to obtain reliable classification performance. The experimental results for the three cooking levels (i.e., undercooked, cooked, and overcooked) showed that six gas sensors conjugated with the CNN algorithm were selected with an F1-measure value above 90%. Meanwhile, two gas sensors coupled with ANN were selected for two cooking levels (i.e., undercooked and cooked), which made it possible to produce a simple electronic nose system. The higher the oven temperature, the faster the electronic nose gave an alarm indicating that the sample had reached the cooked level. In future work, compact and portable equipment is preferred for practical use.

ACKNOWLEDGMENT

REFERENCES
[1] R. Yang and J. Chen, "Mechanistic and machine learning modeling of microwave heating process in domestic ovens: A review," *Foods*, vol. 10, no. 9, p. 2029, Aug. 2021.

[2] A. Adeyinka, O. Olusegun, A. Taiye, L. Mojeed, and O. Heritage, "Development and performance evaluation of dual powered baking oven," *Leonardo Electron. J. Pract. Technol.*, vol. 17, no. 3, pp. 1–15, 2018.

[3] I.-S. Jeong, G.-W. Park, Y.-S. Kim, K.-H. Park, H. Hwangbo, S.-Y. Lee, and S.-G. Kim, "Effects of a novel convection heated roasting method on aroma and umami taste components of vegetable- and meat-based extracts," *J. Food Qual.*, vol. 2022, pp. 1–13, Feb. 2022.

[4] M. Rana, S. Yasmin, M. S. Mahomud, F. Noor, and M. S. H. Sarker, "Development and performance evaluation of an improved electric baking oven for baked products," *Food Sci. Nutrition*, vol. 11, no. 6, pp. 3057–3066, Jun. 2023.

[5] V. Alfeo, E. Bravi, D. Ceccaroni, V. Sileoni, G. Perretti, and O. Marconi, "Effect of baking time and temperature on nutrients and phenolic compounds content of fresh sprouts breadlike product," *Foods*, vol. 9, no. 10, p. 1447, Oct. 2020.

[6] M. M. Z. Azmi, F. S. Taip, S. M. M. Kamal, and N. L. Chin, "Effects of temperature and time on the physical characteristics of moist cakes baked in air fryer," *J. Food Sci. Technol.*, vol. 56, no. 10, pp. 4616–4624, 2019.

[7] E. A. Clark and F. M. Aramouni, "Evaluation of quality parameters in gluten-free bread formulated with breadfruit (*Artocarpus altilis*) flour," *J. Food Qual.*, vol. 2018, pp. 1–12, Sep. 2018.

[8] S. J. Olakanmi, D. S. Jayas, and J. Paliwal, "Applications of imaging systems for the assessment of quality characteristics of bread and other baked goods: A review," *Comprehensive Rev. Food Sci. Food Saf.*, vol. 22, no. 3, pp. 1817–1838, May 2023.

[9] M. Meenu, C. Kurade, B. C. Neelapu, S. Kalra, H. S. Ramaswamy, and Y. Yu, "A concise review on food quality assessment using digital image processing," *Trends Food Sci. Technol.*, vol. 118, pp. 106–124, Dec. 2021.

[10] H. Amani, L. Baranyai, K. Badak-Kerti, and A. M. Khaneghah, "Influence of baking temperature and formulation on physical, sensorial, and morphological properties of pogácsa cake: An image analysis study," *Foods*, vol. 11, no. 3, pp. 1–13, 2022.

[11] N. Ahmed, K. Farooq, S. Shahida, K. U. Haq, S. A. Abbasi, Z. A. Umar, R. Ahmed, and M. A. Baig, "Spectrochemical analysis of Pakistani bakery breads using laser induced breakdown spectroscopy," *Optik*, vol. 226, Jan. 2021, Art. no. 165743.

[12] Y. Wang, J. Zhao, F. Xu, Q. Zhang, Z. Ai, and B. Li, "GC-MS analyses of volatile compounds of steamed breads fermented by Chinese traditional starter 'Jiaozi' from different regions," *J. Food Process. Preservation*, vol. 45, no. 3, pp. 1–13, Mar. 2021.

[13] C. P. Passos, S. Petronilho, A. F. Serôdio, A. C. M. Neto, D. Torres, A. Rudnitskaya, C. Nunes, K. Kukurová, Z. Ciesarová, S. M. Rocha, and M. A. Coimbra, "HS-SPME gas chromatography approach for underivatized acrylamide determination in biscuits," *Foods*, vol. 10, no. 9, p. 2183, Sep. 2021.

[14] Y. Wang, J. Zhao, F. Xu, X. Wu, W. Hu, Y. Chang, L. Zhang, J. Chen, and C. Liu, "GC-MS, GC-O and OAV analyses of key aroma compounds in jiaozi steamed bread," *Grain Oil Sci. Technol.*, vol. 3, no. 1, pp. 9–17, Mar. 2020.

[15] J. Jia, X. Deng, X. Jia, C. Guo, X. Liu, Y. Liu, and X. Duan, "Comparison and evaluation of L. Reuteri and L. rhamnosus-fermented egg yolk on the physicochemical and flavor properties of cookies," *Food Chem., X*, vol. 21, Mar. 2024, Art. no. 101096.

[16] A. Antoniewska, J. Rutkowska, and M. M. Pineda, "Antioxidative, sensory and volatile profiles of cookies enriched with freeze-dried Japanese quince (*Chaenomeles japonica*) fruits," *Food Chem.*, vol. 286, pp. 376–387, Jul. 2019.

[17] R. Yared, "Sensor-based smart oven system to enhance cooking safety," *Proceedings*, vol. 1, no. 2, pp. 1–6, 2016.

[18] M. Misbah, F. Rivai, Z. Kurniawan, D. Muchidin, and D. Aulia, "Identification of diabetes through urine using gas sensor and convolutional neural network," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 1, pp. 520–529, 2022.

[19] D. Aulia, R. Sarno, S. C. Hidayati, and M. Rivai, "Optimization of the electronic nose sensor array for asthma detection based on genetic algorithm," *IEEE Access*, vol. 11, pp. 74924–74935, 2023.

[20] M. Rivai, F. Budiman, D. Purwanto, M. S. A. Al Baid, Tukadi, and D. Aulia, "Discrimination of durian ripeness level using gas sensors and neural network," *Proc. Comput. Sci.*, vol. 197, pp. 677–684, Jan. 2022.

[21] R. Rusinek, M. Gancarz, and A. Nawrocka, "Application of an electronic nose with novel method for generation of smellprints for testing the suitability for consumption of wheat bread during 4-day storage," *LWT*, vol. 117, Jan. 2020, Art. no. 108665.

[22] K. A. Sudama, M. Rivai, D. Aulia, and T. Mujiono, "Electronic nose based on gas sensor array and neural network for indoor hydrogen gas control system," in *Proc. 1st Int. Conf. Inf. Syst. Inf. Technol. (ICISIT)*, Jul. 2022, pp. 187–192.

[23] F. Song, H. Xiang, Z. Li, J. Li, L. Li, and C. Fang Song, "Monitoring the baking quality of tieguanyin via electronic nose combined with GC–MS," *Food Res. Int.*, vol. 165, Mar. 2023, Art. no. 112513.

[24] Radi, M. Rivai, and M. H. Purnomo, "Study on electronic-nose-based quality monitoring system for coffee under roasting," *J. Circuits, Syst. Comput.*, vol. 25, no. 10, Oct. 2016, Art. no. 1650116.

[25] M. Gancarz, U. Malaga-Toboła, A. Oniszczuk, S. Tabor, T. Oniszczuk, M. Gawrysiak-Witulska, and R. Rusinek, "Detection and measurement of aroma compounds with the electronic nose and a novel method for MOS sensor signal analysis during the wheat bread making process," *Food Bioproducts Process.*, vol. 127, pp. 90–98, May 2021.

[26] G. Tucker and F. Gates, *Measuring Oven Exhaust Gases During Bread Baking*. Chipping Campden, U.K.: Campden BRI, 2015, pp. 1–37.

[27] Z. Li, H. Li, Z. Wu, M. Wang, J. Luo, H. Torun, P. Hu, C. Yang, M. Grundmann, X. Liu, and Y. Fu, "Advances in designs and mechanisms of semiconducting metal oxide nanostructures for high-precision gas sensors operated at room temperature," *Mater. Horizons*, vol. 6, no. 3, pp. 470–506, 2019.

[28] A. Ponzoni, C. Baratto, N. Cattabiani, M. Falasconi, V. Galstyan, E. Nunez-Carmona, F. Rigoni, V. Sberveglieri, G. Zambotti, and D. Zappa, "Metal oxide gas sensors, a survey of selectivity issues addressed at the SENSOR lab, Brescia (Italy)," *Sensors*, vol. 17, no. 4, p. 714, Mar. 2017.

[29] Figaro. (2013). *TGS 2602 for the Detection of Air Contaminants*. FIGARO Engineering Inc. [Online]. Available: http://www.figarosensor.com/products/docs/TGS2602-B00 (0913).pdf

[30] D. Cabrera, F. Sancho, J. Long, R.-V. Sánchez, S. Zhang, M. Cerrada, and C. Li, "Generative adversarial networks selection approach for extremely imbalanced fault diagnosis of reciprocating machinery," *IEEE Access*, vol. 7, pp. 70643–70653, 2019.

[31] G.-H. Shin and J.-W. Hur, "Correlation coefficient based optimal vibration sensor placement and number," *Sensors*, vol. 22, no. 3, p. 1207, Feb. 2022.

[32] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, "Feature selection–based detection of covert cyber deception assaults in smart grid communications networks using machine learning," *IEEE Access*, vol. 6, pp. 27518–27529, 2018.

[33] P. Giungato, G. de Gennaro, P. Barbieri, S. Briguglio, M. Amodio, L. de Gennaro, and F. Lasigna, "Improving recognition of odors in a waste management plant by using electronic noses with different technologies, gas chromatography–mass spectrometry/olfactometry and dynamic olfactometry," *J. Cleaner Prod.*, vol. 133, pp. 1395–1402, Oct. 2016.

[34] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE Access*, vol. 6, pp. 33789–33795, 2018.

[35] V. P. Kour and S. Arora, "Particle swarm optimization based support vector machine (P-SVM) for the segmentation and classification of plants," *IEEE Access*, vol. 7, pp. 29374–29385, 2019.

[36] C. Ding, T.-Y. Bao, and H.-L. Huang, "Quantum-inspired support vector machine," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7210–7222, Dec. 2022.

[37] C. Zou, A. Müller, U. Wolfgang, D. Rückert, P. Müller, M. Becker, A. Steger, and E. Martens, "Heartbeat classification by random forest with a novel context feature: A segment label," *IEEE J. Transl. Eng. Health Med.*, vol. 10, pp. 1–8, 2022.

[38] V. Jain, A. Phophalia, and S. K. Mitra, "HML-RF: Hybrid multi-label random forest," *IEEE Access*, vol. 10, pp. 22902–22914, 2022.

[39] H. Soliman, "Random forest based searching approach for RDF," *IEEE Access*, vol. 8, pp. 50367–50376, 2020.

[40] M. Gencturk, A. A. Sinaci, and N. K. Cicekli, "BOFRF: A novel boosting-based federated random forest algorithm on horizontally partitioned data," *IEEE Access*, vol. 10, pp. 89835–89851, 2022.

[41] S.-E. Ryu, D.-H. Shin, and K. Chung, "Prediction model of dementia risk based on XGBoost using derived variable extraction and hyper parameter optimization," *IEEE Access*, vol. 8, pp. 177708–177720, 2020.

[42] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A pedestrian detection method based on genetic algorithm for optimize XGBoost training parameters," *IEEE Access*, vol. 7, pp. 118310–118321, 2019.

[43] M. Chen, Q. Liu, S. Chen, Y. Liu, C.-H. Zhang, and R. Liu, "XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system," *IEEE Access*, vol. 7, pp. 13149–13158, 2019.

[44] A. Zaib, T. Ballal, S. Khattak, and T. Y. Al-Naffouri, "A doubly regularized linear discriminant analysis classifier with automatic parameter selection," *IEEE Access*, vol. 9, pp. 51343–51354, 2021.

[45] H. Zhao, Z. Wang, and F. Nie, "A new formulation of linear discriminant analysis for robust dimensionality reduction," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 4, pp. 629–640, Apr. 2019.

[46] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, A. M. Umar, O. U. Linus, H. Arshad, A. A. Kazaure, U. Gana, and M. U. Kiru, "Comprehensive review of artificial neural network applications to pattern recognition," *IEEE Access*, vol. 7, pp. 158820–158846, 2019.

[47] G. Zhang, V. Davoodnia, A. Sepas-Moghaddam, Y. Zhang, and A. Etemad, "Classification of hand movements from EEG using a deep attention-based LSTM network," *IEEE Sensors J.*, vol. 20, no. 6, pp. 3113–3122, Mar. 2020.

[48] Z. Hu, L. Chen, Y. Luo, and J. Zhou, "EEG-based emotion recognition using convolutional recurrent neural network with multi-head self-attention," *Appl. Sci.*, vol. 12, no. 21, p. 11255, Nov. 2022.

[49] Y. A. M. Alsumaidaee, C. T. Yaw, S. P. Koh, S. K. Tiong, C. P. Chen, T. Yusaf, A. N. Abdalla, K. Ali, and A. A. Raj, "Detection of corona faults in switchgear by using 1D-CNN, LSTM, and 1D-CNN-LSTM methods," *Sensors*, vol. 23, no. 6, p. 3108, Mar. 2023.

[50] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019.

[51] J. A. Célia, O. Resende, M. S. D. Lima, J. S. Correia, K. B. D. Oliveira, and K. P. Takeuchi, "Technological properties of gluten-free biscuits from sorghum flour granifero (*Sorghum bicolor* (L.) Moench)," *Food Sci. Technol.*, vol. 42, pp. 1–7, 2022.

MUHAMMAD RIVAI (Member, IEEE) received the Ph.D. degree, in 2006. He is currently a Lecturer with the Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. His research interests include sensors, chemical sensors, gas sensors, electronics, robotics, and artificial intelligence applications.

DAVA AULIA is currently pursuing the master's degree with the Informatics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. His research interests include artificial intelligence, deep learning, the Internet of Things (IoT), embedded systems, and gas sensors.

● ● ●