Computers and Electronics in Agriculture 195 (2022) 106838

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

# The effect of gas concentration on detection and classification of beef and pork mixtures using E-nose

Sulaiman Wakhid [a], Riyanarto Sarno [b,*], Shoffi Izza Sabilla [b]

[a] Departement of Information Technology Management, Faculty of Creative Design and Digital Business, Institut Teknologi Sepuluh Nopember (ITS) Sukolilo, Surabaya 60111, Indonesia
[b] Department of Informatics, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember (ITS) Sukolilo, Surabaya 60111, Indonesia

ARTICLE INFO

*Keywords:*
Beef
Pork
E-Nose
Gas concentration
Classification
Machine learning

ABSTRACT

Examining the purity of meat is a classical problem in developing countries, especially in Indonesia. The high economic value of beef causes counterfeiting to occur frequently. The forgery process is done through the simple practice of mixing in a certain percentage of pork. Several recent studies have shown that e-noses can examine beef purity through gas detection. This study aimed to determine the effect of gas concentration on the results of detection and classification of beef and pork mixtures by characterizing different meat samples in 3 chambers with a different size. The meat mixture dataset was retrieved by an e-nose device with an array of MQ series sensors that are sensitive to chemical scents. Classification of the meat mixtures was done in several stages: data acquisition from the 3 different sample chambers, statistical feature extraction, classification, ensemble learning, and performance evaluation based on a confusion matrix. The experimental results from this study indicate that the sample chamber with the highest gas concentration yielded the highest accuracy. The best accuracy result, i.e. 95.71%, was obtained with a 50-ml sample chamber using an ensemble method with the statistical parameters of kurtosis and skewness.

## 1. Introduction

Beef is among the food ingredients with the highest consumption levels in the world. Moreover, the Food and Agriculture Organization (FAO) has predicted that beef will remain a main food source with high consumption demand until 2050 (Wijaya et al., 2021). According to a survey conducted by the Central Bureau of Statistics of Indonesia (*Badan Pusat Statistik Indonesia*) in 2019, the demand for beef consumption in Indonesia continues to increase. The majority of Indonesia's population is Muslim, making the demand increase in certain seasons, such as during the fasting month and Eid al-Fitr. This creates a gap between the production and consumption of beef, which allows its price in the market to soar in a short time. This makes beef a commodity with a high economic value (Rustinsyah, 2019), which creates an opportunity for some beef sellers and suppliers to commit counterfeiting fraud by mixing in a percentage of pork. In Muslim religious regulations, consuming any amount of pork is prohibited. This issue has a direct impact on buyers, especially Muslims (Nakyinsige et al., 2012).

Beef that has been mixed with pork is very difficult to distinguish by the human senses. It takes special techniques and expertise to analyze meat mixtures, visually (texture and color), physically (tenderness, aroma, and taste), and chemically (compounds) or biologically (microorganisms) (Wijaya et al., 2017; Liu et al., 2020). Beef and pork each have their own distinctive aroma compounds. Thus it is possible to detect them by using an e-nose (electronic nose) device (Alim et al., 2019; Wijaya et al., 2019; Shao et al., 2018). The e-nose sensor circuit consists of electrochemical sensors that are sensitive to chemical compounds. The smell from a mixture of beef and pork is captured by the e-nose to serve as ground-truth data (Wijaya et al., 2017). Several studies have shown that e-noses can recognize pure beef and pure pork by using

*Abbreviations:* E-nose, Electronic Nose; FAO, Food and Agriculture Organization; SVM, Support Vector Machine; LR, Logistic Regression; DTC, Decision Tree Classifier; ANN, Artificial Neural Network; BPNN, Back Propagation Neural Network; PPM, Parts Per Million; MOS, Metal Oxide Semiconductor; AVG, Average; STD, Standatd Deviation; KRT, Kurtosis; SKW, Skewness; MIN, Minimum; MAX, Maximum; RBF, Radial Basis Function; MLP, Multilayer Perceptron; RS, Sensor Resistance; CO, Carbon Monoxide; TP, True Positives; TN, True Negatives; FN, False Negatives; FP, False Positives.

* Corresponding author.
*E-mail address:* riyanarto@if.its.ac.id (R. Sarno).

https://doi.org/10.1016/j.compag.2022.106838
Received 20 August 2021; Received in revised form 25 February 2022; Accepted 28 February 2022
Available online 9 March 2022

a naive Bayes classification method (Wijaya et al., 2017). Another study was carried out to identify mixtures of beef and pork, using as many as 5 data classes (Laga and Sarno, 2020). In addition to objects in the form of meat, other studies have investigated the capabilities of e-noses in detecting other objects, such as the aroma of civet coffee and the sweetness of pineapples (Wijaya et al., 2017; Harsono et al., 2020; Hasan et al., 2020). Hence, e-noses have the capability to recognize the aroma of various objects (including meat) through aromatic gases.

One of the components of an e-nose is the sample chamber to accommodate the aromatic gas (Zhiyi et al., 2017; Li et al., 2016). The size of the sample chamber affects the concentration of the gas. The main objective of this study was to determine the effect of gas concentration on the results of detection and classification of mixed beef and pork using an e-nose. Detection and classification were carried out in 3 sample chambers of different sizes with 7 classes of meat mixtures. The sample chambers used were beaker glasses with a size of 50 ml, 150 ml, and 250 ml. The meat samples were divided into 7 percentage classes. The first class was pure beef at a rate of 100%. The second, third, fourth, fifth, and sixth classes were a combination of beef mixed with pork at a rate of 10%, 25%, 50%, 75%, and 90%. The last class was pure pork at a rate of 100%. The beef and pork were purchased on the same day and at the same shop. The livestock had been slaughtered on the same day as the day of purchase. The aim of using fresh meat is to avoid aroma dissimilarity caused by oxidation of the chemical aroma of rotting meat (Tian et al., 2013).

The classification process carried out in this study started with data acquisition. The gas data was obtained using an e-nose device as ground-truth data. Ground-truth data is raw data that must be pre-processed to eliminate unnecessary data. Furthermore, feature extraction was carried out using the following statistical parameters: standard deviation, mean, kurtosis, skewness, minimum, and maximum value. A comparative analysis between the statistical parameters and several classification methods was carried out to obtain the best combination. Four classification methods based on machine learning were used, namely support vector machine (SVM), logistic regression (LR), decision tree classifier (DTC), and artificial neural network (ANN). Each of these methods produces a model for the ensemble method to improve the accuracy results. Lastly, the accuracy rate was obtained from evaluation by k-fold cross-validation and a confusion matrix (Wong, May 2017).

This paper is organized into several sections: Sections 1 and 2 describe the research background and related work. Section 3 describes the materials and methods proposed for solving the problems. Sections 4 describe the evaluation and the results obtained. The last section presents the conclusion. There are three main contributions of this research, i.e. (i) knowing the optimum combination of statistical parameters for the classification of beef and pork mixtures; (ii) detecting 7 classes of beef and pork mixtures with 3 different sample chamber sizes and determining the classification accuracy; (iii) determining the effect of gas concentration on the results of detection and classification of the beef and pork mixtures.

## 2. Related work

Previous researchers have conducted studies related to e-noses with various kinds of experimental objects. E-noses have not only been used to detect meat but also for example to detect civet and non-civet coffee (Wakhid et al., 2020). In this study, experiments were carried out with 6 classes of coffee data taken from 3 different regions in Indonesia. Aroma detection was carried out with a series of MQ sensors connected to an Arduino board. Using the best combination of statistical parameters determined through average and standard deviation, the accuracy rate was 97%. An e-nose has also been used to recognize the level of sweetness of pineapples (Ananas comosus (L.) Merr.) (Hasan et al., 2020). In this study, to determine the level of ripeness of pineapples, the level of maturity was measured using a Brix meter with 3 intervals of maturity levels: low, medium, and high. Signal processing methods such as PCA and the mother wavelet were used to optimize the ground-truth data. The highest classification accuracy was obtained by the KNN method, with a value of 82%.

Research using meat has been conducted to detect the freshness of meat using an e-nose assembled using a Raspberry Pi with a gas sensor and a color sensor (Rivai et al., 2018). The combination of odor and visual information was used in this study to obtain different patterns for three categories of freshness. The result was an e-nose that was able to distinguish fresh and rotten meat at an accuracy rate of 100%. Meanwhile, the difference between fresh and half-rotten meat was distinguished with an accuracy rate of 80%. Another study tried to prevent counterfeiting lamb with pork through e-nose scent detection (Tian et al., 2013). Detection was done through a combination of traditional methods (pH and color evaluation) with an e-nose. The best results were obtained with a back propagation neural network (BPNN) model. Another study examined the effect of temperature on classification with 5 classes of beef and pork mixtures using an e-nose (Laga and Sarno, 2020). The results showed the effect of temperature on classification accuracy, which was highest at a temperature of –22 °C.

Previous studies have shown that e-noses have the capability to detect organic compounds, including meat. However, no research has been done yet on the effect of gas concentration on the accuracy of the classification of the purity of beef, specifically the effect of using sample chambers of different sizes. One previous study only distinguished pure beef and pork, with an accuracy of only 75% (Wijaya et al., 2017). Meanwhile, another study only attempted to find the effect of temperature on the classification accuracy of beef purity (Laga and Sarno, 2019, 2019). These studies did not discuss the sample chamber used. For this reason, the present study examined the effect of gas concentration on the detection and classification of beef purity through the use of 3 sample chambers of different sizes.

## 3. Proposed materials and method

This research focused on analyzing the effect of different gas concentrations of beef and pork on the classification results. Several steps are proposed to obtain the classification accuracy rates from three sample chambers. The classification scheme used in this study is shown in Fig. 1. Raw data was obtained from the e-nose's sensors (Wang et al., 2020). This data was in the form of a digital signal that had been converted to ppm (parts per million), after which feature extraction was carried out to obtain the statistical parameter values. To obtain predictions, the data was processed using several machine learning classification algorithms, i.e. support vector machine (SVM), logistic regression (LR), decision tree classifier (DTC), and artificial neural network (ANN).

The classification results obtained were evaluated to determine the performance and accuracy of each algorithm. This process was repeated for each combination of statistical parameters and features used to find the best performing combination. To get the optimal accuracy, ensemble calculation was carried out using voting. The calculation was carried out on the data obtained from all three sample chambers. Finally, evaluation was done by comparing the accuracy rates based on the sample chamber used.

### 3.1. E-nose system and data collection

The device used in this study was a vacuum type e-nose with gas scent suction. The e-nose was built using an Arduino Mega 256 board with a sensor array inside, as shown in Fig. 2. The sensor array is constructed using Metal oxide semiconductor (MOS) gas sensor type also known as chemiresistors because the detection proccess is based on the change in the resistance of the sensing material when the gas is in contact with the material. The MOS gas sensor was chosen because the advantages of low-cost, fast response, and sensitivity (Patel, 2014). In this experiment using nine MOS gas sensors from MQ family. Each

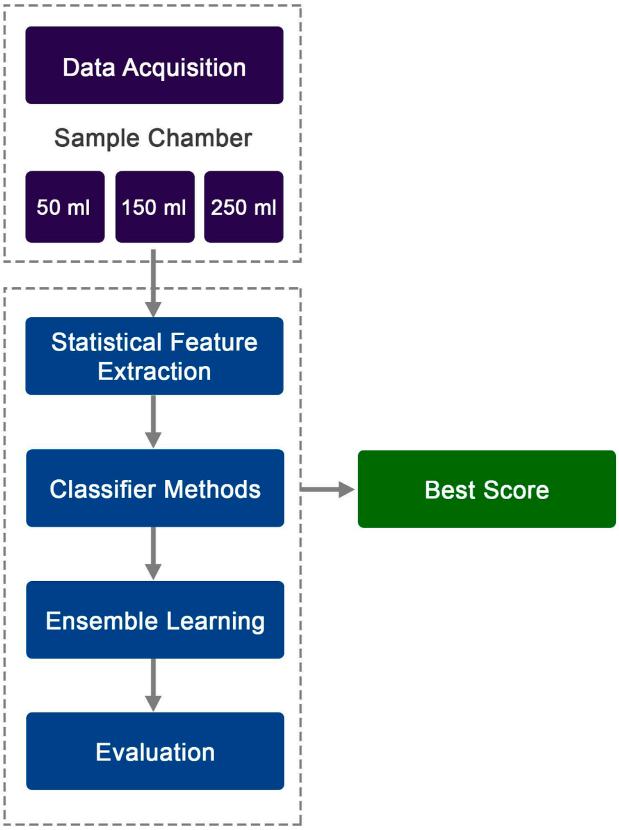


**Fig. 1.** The process of obtaining classification scores from the 3 different sample chambers.

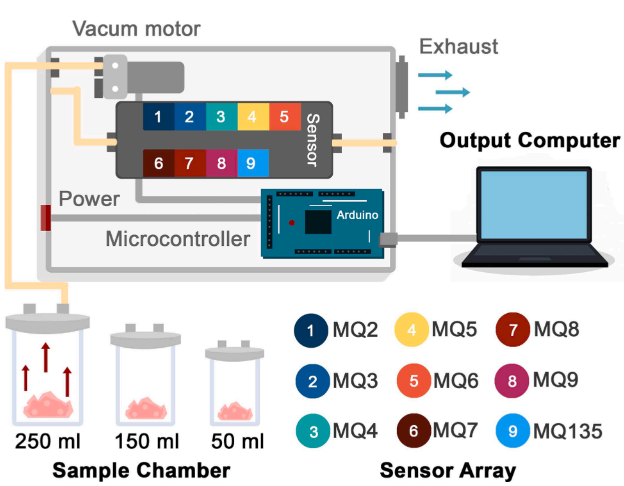


**Fig. 2.** E-Nose system design with MQ family sensors.

sensor has sensitivity toward specific gas compounds (Gu et al., 2013), as shown in Table 1.

The MQ gas sensor has been widely applied in various e-nose studies including the detection of civet coffee aroma (Wakhid et al., 2020), detection of the sweetness of pineapple (Hasan et al., 2020), halal authentication (Wijaya et al., 2017), and monitoring of meat quality (Wijaya et al., 2021; Jiang et al., 2015). The selectivity of the gas sensor is determined based on the gas produced by the meat protein degradation process which includes carbon dioxide gas, hydrogen sulfide, ammonia, methane (Dent et al., 2004; Zheng et al., 2019). As well as other compounds aldehydes, ketones, hydrogen, and alcohol were also seen from the degradation of meat carbohydrates (Hui, 2016; Han et al., 2015).

**Table 1**
MQ Family Sensors and Measured Gas Contents used in the E-Nose System.

| Sensor | Gas Contents |
|---|---|
| MQ 2 | $CH_4$, hydrogen ($H_2$), carbon monoxide (CO), LPG, alcohol, propane, air |
| MQ 3 | Carbon monoxide (CO), benzene, alcohol, methane, hexane, LPG, air, $CH_4$ |
| MQ 4 | $CH_4$, $H_2$, alcohol, smoke |
| MQ 5 | $CH_4$, $H_2$, alcohol, LPG, carbon monoxide (CO), air |
| MQ 6 | LPG, butane |
| MQ 7 | Carbon monoxide (CO) |
| MQ 8 | Hydrogen ($H_2$) |
| MQ 9 | Carbon monoxide (CO), flammable gases |
| MQ 135 | $NH_3$, mono nitrogen ($NO_{x)}$), alcohol, benzene, smoke, $CO_2$ |

Data collection through the e-nose system involves 4 processes that must be executed sequentially: warming, flushing, sensing, and cleaning (Sabilla et al., 2020), each process taking a certain amount of time. Starting from turning on the e-nose, the warming-up process takes 5 min, the flushing process takes 1 min, sensing takes 2 min, and cleaning takes 3 min. During the flushing process, the aroma gas from each sample chamber is sucked in and flowed into the sensor chamber. The captured gas is processed and converted from a stationary signal to a digital signal by the Arduino. Thus, raw data is obtained in the form of numeric values for each sensor.

### 3.2. Statistical feature extraction

The gas produced by the chemical processes of the meat is captured by the sensor array of the e-nose system. The gas is a stationary signal so that statistical parameter values can be obtained (Hariyanto et al., 2018). Statistical parameters are measurements of the characteristics of an object. In this case, the statistical parameters provide a unique value for the aroma of each meat mixture. In this study, six statistical parameters were used: average (AVG), standard deviation (STD), kurtosis (KRT), skewness (SKW), minimum (MIN), and maximum (MAX). The value of each statistical parameter can be obtained by the six equations below:

$$Average\ \ (\overline{x}) = \frac{1}{n}\left(\sum\nolimits_{i=1}^{n} x_i\right) \tag{1}$$

$$Standard\ \ Deviation\ \ (\sigma) = \sqrt{\frac{\sum_{i=1}^{n}(xi-\mu)^2}{n-1}} \tag{2}$$

$$Kurtosis = \frac{1}{T\sigma^4}\sum_{t=1}^{T}(T_t-\mu)^4 \tag{3}$$

$$Skewnes = \frac{1}{T\sigma^3}\sum_{t=1}^{T}(r_t-\mu)^3 \tag{4}$$

$$Maximum = max(x_i) \tag{5}$$

$$Minimum = min(x_i) \tag{6}$$

The use of appropriate statistical parameters can affect the accuracy of the results (Wakhid et al., 2020). For this reason, experiments with different combinations of statistical parameters were carried out. Six different combinations were tried, as shown in Table 2. For each combination the accuracy was calculated so that a comparison could be made (Sarno et al., 2020).

### 3.3. Classification methods

The classification stage is one of the main topics of this study. A comparison between 4 different classification methods was carried out. Classification is a branch of machine learning that falls into the category

**Table 2**
Statistical Parameter Combinations.

| Combination | Parameters |
|---|---|
| 1 | STD-AVG |
| 2 | STD-AVG-MIN-MAX |
| 3 | STD-AVG-KRT-SKW |
| 4 | STD-AVG-KRT-SKW-MIN-MAX |
| 5 | KRT-SKW |
| 6 | KRT-SKW-MIN-MAX |

of supervised learning. Various classification methods can be used (Ahmed et al., 2021). The classification methods used in this study were: support vector machine (SVM), logistic regression (LR), decision tree classifier (DTC), and artificial neural network (ANN).

SVM is a classification model that uses the technique of finding the best hyperplane by maximizing the distance between classes. A hyperplane is a function that can be used to separate between two classes in the input space. The best separating hyperplane between the two classes can be found by measuring the margin distance from the hyperplane and then finding the maximum point (Phillips and Abdulla, 2021). The margin is the distance between the hyperplane and the closest pattern from each existing class. The margin is calculated by Eq. (7):

$$\left(\frac{w}{||w||}(x_1 - x_2)\right) = \frac{w}{||w||} \tag{7}$$

SVM can be used flexibly in the classification of linear and non-linear data. Non-linear data can be classified using kernel functions, which try to map the data to higher dimensions for better accuracy. Examples of kernel functions are: the polynomial kernel, the radial basis function (RBF) kernel, and the sigmoid kernel (Mohammadi et al., 2021). The machine learning model implementation used in this study came from the Sklearn library (Pedregosa et al., 2011). The parameters used were 'kernel' = rbf and 'random_state' = 8.

Logistic regression is a classification model used to predict whether something is true or false (0 or 1). In contrast to linear regression, where the line formed is straight, in logistic regression the line formed between points 0 to 1 is similar to the letter S, as shown in Fig. 3, where the relationship between the target variable and the input variable used by the logistic function is measured. One advantage of LR is its ability to describe the degree of the relationship between the dependent variable and several independent variables. The main advantage of LR is that it does not require a value with a normal distribution so that the causal variable can be a continuous value, a discrete value, or a combination of both (Ghosh and Dey, 2021). The logistic regression used in this study was multinomial, or multilabel, which has 2 or more outputs (Książek et al., 2021). The parameters used in logistic regression were 'max_iteration' = 1000, 'random_state' = 8, and 'solver' = lbfgs.

DTC is a classification model using a tree, or hierarchical, structure. The concept of a decision tree is to convert data into a decision tree and decision rules (Cervantes et al., 2018). It consists of a root, several branches, several nodes, and a number of leaves. Fig. 4 shows the structure of a simple binary tree with a top-down structure. It is known that the value of $X = (X_1, X_2, X_3)$ is a vector of the independent variables, while the threshold and leaf values are indicated by $T_i$ and $L_i$. The iteration of developing the tree structure continues until the stopping criterium or criteria are met. DTC can be used in solving both classification and regression problems (Ghiasi et al., 2020). The main benefit of the DTC method is its ability to break down complex decision-making processes into simpler ones so that decision-makers are better able to interpret solutions to problems. In this study, the parameters used in the application of the DTC method were 'random_state' = 8 and 'criterion' = gini.

ANN is an information processing method with an approach that is similar to how a biological nervous system works, specifically human brain cells, when processing information (Goudarzi et al., 2021). An ANN consists of a large number of information processing elements (neurons) that are interconnected and work together in solving a particular problem, where the problem can be a classification or a prediction problem. In a neuron network there are several layers, namely an input layer, a hidden layer, and an output layer (Ahmad et al., 2017). The input layer serves to facilitate the convergence of the neural network. In the hidden layer (main neural layer), the variable input data is collected, while the output layer works to collect the resulted values.

$$Sigmoidal : f(x) = \frac{1}{1 + exp^{-x}} \tag{8}$$

$$Tansig : f(x) = \frac{2}{2 + exp^{-2x}} \tag{9}$$

In this study, a multilayer perceptron (MLP) structure was used for classification. MLP can be used mathematically to perform stochastic estimation of multivariate functions represented by Eqs. (8) and (9). The implementation of ANN in this study used the parameters 'hidden_ layer' = 10, 'max_iteration' = 1000, and 'random_state' = 8.

### 3.4. Ensemble method

Ensemble learning is an approach that can help improve the prediction accuracy of machine learning (Raza, 2019). This method uses a meta-algorithm that combines several machine learning techniques into one prediction model (Li and Lin, 2021). The ensemble method in this study used voting. Voting (or stacking) is an ensemble model in which a new model is trained from the combined predictions of two or more

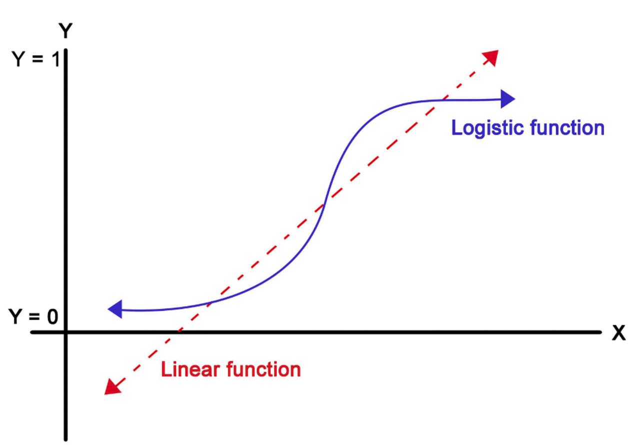


**Fig. 3.** Logistic regression curve.

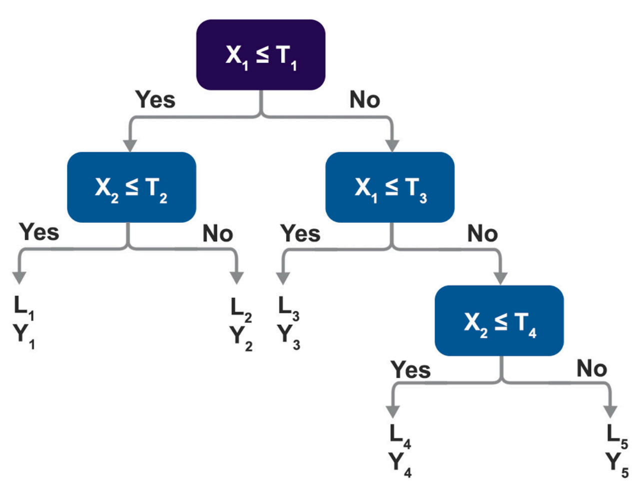


**Fig. 4.** Top-down binary tree structure.

previous models. The concept of merging models is called meta-learning (Kumari et al., 2021). This method can combine models from different classification types. This is possible because the first-level learner uses the original training data set. Then a new data set is generated to be trained by the second-level learner, where the output of the first-level learner is considered as the original temporary input feature. An illustration of the voting method is shown in Fig. 5.

The voting approach has two types of voting mechanisms that can be done, namely soft voting and hard voting (Malikhah et al., 2021). In the type of soft voting the probability vectors for each class that are predicted for each classifier are summed and averaged. The final prediction class label taken is the one that corresponds to the highest value obtained. The voting method approach in this study uses the hard voting type where the final prediction result is the class that gets more than half of the votes or the majority vote. The final prediction of the most votes is obtained using Eq. (10), where $H(x)$ is the final prediction of the most votes; $T$ is the base of the classifier number;*arg* is used to select the class with the most votes that gets more than half of the votes; and $h_i^y(x)$ is the prediction for each base classifier.

$$H(x) = \underset{y \in Y}{argmax} \sum\nolimits_{i=1}^{T} h_i^y(x) \tag{10}$$

## 4. Results and discussion

### 4.1. Data acquisition

This study used seven ground-truth data for beef mixed with 7 different percentages of pork. Sampling was carried out in 3 sample chambers with different sizes: 50 ml, 150 ml, and 250 ml. Thus, the total number of classes obtained from data collection was 21. Each class was given a unique label for easy identification, as shown in Table 3.

Data was collected 50 times for each class at room temperature between 25 °C and 30 °C (Zhou et al., May 2021; Shen et al., 2020). Each meat sample weighed 10 g (Zheng et al., 2019). The meat was a mixture of beef and pork that had been roughly chopped. The meat was bought from traditional traders who cut it every day so that the meat used was fresh.

The detection of meat aroma gas concentration was carried out using the e-nose device described above by inserting a sample into the chamber. The aroma of the meat sample was captured through one MQ 7 sensor installed in the e-nose. The data was transmitted to the Arduino microcontroller. The final output value of the e-nose was a gas concentration value in ppm (parts per million). In this study, the gas concentrations in the three sample chambers of different sizes were compared, as shown in Fig. 6. Based on the graph it can be seen that the sensor resistance value (RS) was lowest in the 50-ml sample chamber. This shows that the gas concentration was highest in the smallest sample chamber.

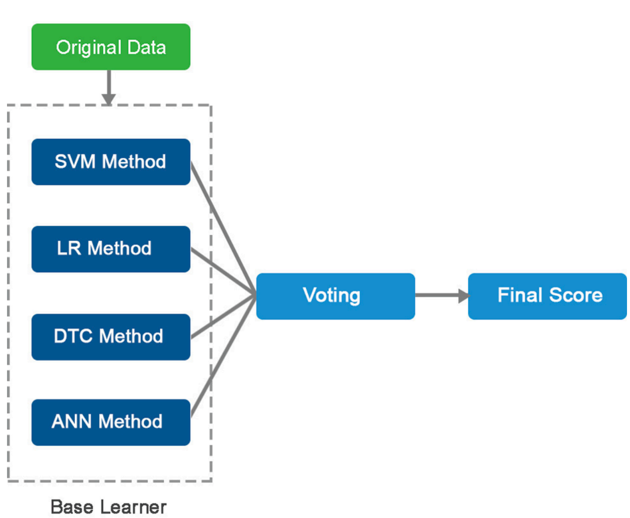


**Fig. 5.** Ensemble method with voting approach.

**Table 3**
Meat Mixture Data Labels based on Sample Chamber.

| Data Label | | | Meat Percentage |
|---|---|---|---|
| 50 ml | 150 ml | 250 ml | |
| AS000 | BS000 | CS000 | Beef 100% Pork 0% |
| AS010 | BS010 | CS010 | Beef 90% Pork 10% |
| AS025 | BS025 | CS025 | Beef 75% Pork 25% |
| AS050 | BS050 | CS050 | Beef 50% Pork 50% |
| AS075 | BS075 | CS075 | Beef 25% Pork 75% |
| AS090 | BS090 | CS090 | Beef 10% Pork 90% |
| AS100 | BS100 | CS100 | Beef 0% Pork 100% |

### 4.2. Combination of statistical parameter extraction

The e-nose device begins to capture digital signals after the heating process, namely at the flushing, sensing, and cleaning stages. These three processes produce a distribution signal with a pattern as shown in Fig. 7.

In the sensing process, the rise in gas value of each class is different and contains unique features that represent the characteristics of the class concerned. At another stage, it is a sloping graph where each class has almost the same value so that the statistical parameter extraction only uses the data from the sensing process.

Statistical parameter extraction is carried out to obtain the characteristic value of each gas content produced by the sensor responese (Sarno et al., 2020). Extraction of gas characteristics in this study used the following six statistical parameters: average (AVG), standard deviation (STD), kurtosis (KRT), skewness (SKW), minimum (MIN), and maximum (MAX). To obtain the average, the gas content values of each sample are summed and divided by the total amount of data generated per sampling $n$, as stated in Eq. (1). The standard deviation parameter can be calculated through Eq. (2). This parameter is used to determine the characteristics of the gas through the variance distribution of sample data values. If the value of $\sigma$ is high, it indicates that the data value spreads from the middle value of $\bar{x}$. Skewness is degree asymmetry of a data distribution. To get the skewness can be done with the formula in Eq. (3) where $t$ is the value of each observed data, $\sigma$ is the standard deviation of the sample data, is the average of the sample data, and $T$ is the number of observations that have been made. Kurtosis is degree of

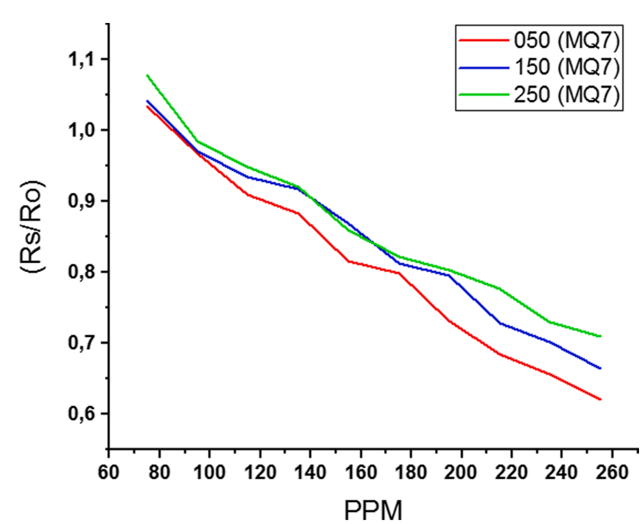


**Fig. 6.** Comparison of gas concentrations by the MQ 7 sensor in the 3 different sample chambers.

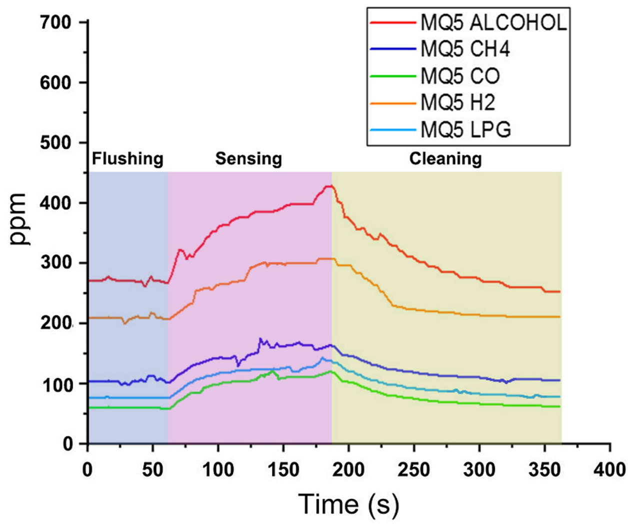


**Fig. 7.** Digital signal pattern from the MQ5 sensor.

tailedness a distribution data that measured relative to the normal distribution. Hence the greater the value of kurtosis, the curve will be tapper. Kurtosis formula can be expressed in Eq. (4). The maximum parameter value is the highest value of the data distribution while the minimum parameter value is the lowest value, as stated in the Eq. (5) and Eq. (6).

The sensor array uses 8 gas sensors as shown in Table 1. Each sensor has gas content that can be extracted. The total amount of gas content that can be extracted from 8 sensors is 48 gas content. Each value of each gas content is extracted using 6 statistical parameters so that there are a total of 288 features. Table 4 is the result of 6 statistical parameter features generated by the MQ 5 sensor for carbon monoxide (CO) gas content. One row of feature extraction results represents one sample of the data obtained, so that for each sample chamber (7 classes) there are 350 rows of data lines. To get the optimal accuracy rate, this study made six trial combination of statistical parameters as shown in Table 2. The first experiment using STD-AVG is combining the standard deviation with the average. The second is using STD-AVG-MIN-MAX by adding the maximum minimum parameter value. The three STD-AVG-KRT-SKW replace the maximum minimum value with kutosis skewness. The four STD-AVG-KRT-SKW-MIN-MAX are combining all parameter values. The five KRT-SKW only use kurtosis and skewness and the last one is KRT-SKW-MIN-MAX.

### 4.3. Classification experiment results

Classification was done using 4 machine learning methods, i.e. SVM, LR, DTC, and ANN. Calculations were made for each of the three sample chambers for 7 classes. The experiment was carried out in 3 steps: first, classification of the data obtained from the 50-ml sample chamber; second, classification of the data obtained from 150-ml sample chamber; and third, classification of the data obtained from the 250-ml sample chamber. The data used was statistical parameter data that had passed the feature extraction process. The classification model evaluation technique used in this experiment is k-fold cross-validation. Using a number of k = 10 the data is divided into 10 folds of approximately equal size, so there are 10 subsets of data to evaluate the performance of the model or algorithm. For each of the 10 data subsets, CV will use 9 folds for training and 1 testing. (Jiang and Wang, 2017). The process of evaluating the results of the classification method was done using a visual tool in the form of a confusion matrix (Düntsch and Gediga, 2020).

The confusion matrix contained cases that were classified correctly and cases that were classified incorrectly. In this case, 4 values were obtained: true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP) (Luque et al., 2019). Fig. 8 shows the confusion matrix from the classification results of the SVM method on the 50 ml-sample chamber. These results show that the highest prediction value was obtained for classes SA000, SA010, and SA025, with 48 true positives. The lowest prediction value was obtained for classes AS050 and AS075, with 46 true positives.

Through the results of the confusion matrix, the performance evaluation values of the output metrics were obtained in the form of precision, recall, f1-score, and accuracy. The values were obtained with the following equations:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{11}$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{12}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{13}$$

$$F1Score = \frac{2 \times precision \times recall}{precision + recall} \times 100\% \tag{14}$$

The accuracy rate is a percentage based on the level of similarity between the predicted value and the actual value. Precision is the percentage of the number of selected attributes that are relevant. The recall value, or sensitivity, is the ratio of true positive predictions compared to the overall data that are true positive. The F1 score is the average of precision and sensitivity (Devi et al., 2019).

**Table 4**
Results of Statistical Parameter Extraction.

| No. | Class | Sensors | | | | | |
|---|---|---|---|---|---|---|---|
| | | AVG MQ5 (CO) | STD MQ5 (CO) | KRT MQ5 (CO) | SKW MQ5 (CO) | MIN MQ5 (CO) | MAX MQ5 (CO) |
| 1 | AS000 | 8.69 | 11.87 | 0.17 | −2.45 | 1.34 | 12.46 |
| 2 | AS000 | 8.70 | 11.45 | 0.17 | −2.40 | 1.32 | 12.35 |
| … | … | … | … | … | … | … | … |
| 175 | AS050 | 12.58 | 8.76 | −1.29 | 0.38 | 5.77 | 16.31 |
| 176 | AS050 | 12.59 | 8.21 | −1.30 | 0.38 | 5.43 | 16.31 |
| … | … | … | … | … | … | … | … |
| 349 | AS100 | 15.09 | 5.77 | −2.22 | 0.25 | 9.23 | 19.17 |
| 350 | AS100 | 14.12 | 5.77 | −1.98 | 0.27 | 9.03 | 19.17 |

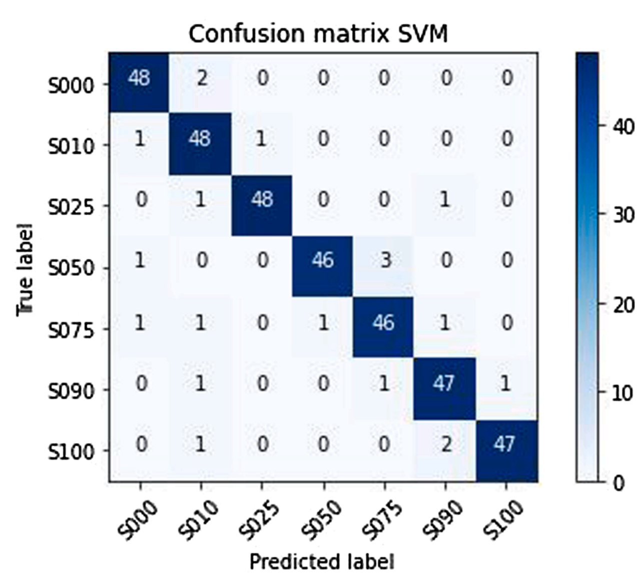


**Fig. 8.** Confusion matrix from the SVM method used on the 50-ml sample chamber.

Table 5 shows the result of the performance evaluation of the SVM method using the 50-ml sample chamber data. It can be seen that the average precision, recall, and f1 scores were above 90% except for the precision value for the AS010 class, which was 89%. The final result of the accuracy obtained was 94.28%.

### 4.4. Ensemble and evaluation performance

The accuracy rate was calculated for each of the four machine learning methods used. Table 6 shows the accuracy rates for the 50-ml sample chamber. The highest accuracy rate was obtained using SVM (94.57%), followed by LR (93.71%), ANN (92.85%), and DTC (91.14%). The accuracy results could still be improved using the ensemble method. The ensemble approach used in this study was a voting technique. The four classification methods were combined by a meta-learner algorithm. Thus, an accuracy rate of 95.71% could be obtained, i.e. an increase of 1.14%.

Further optimization of the accuracy rate was carried out by combining the statistical parameters and the methods as shown in Table 2. This process was carried out on the data from each of the 3 sample chambers. Table 7 is the result of the overall recapitulation of the accuracy rates from all experiments. The table shows that the different combinations produced different results. In the 50-ml sample chamber experiment, the highest accuracy rate was obtained from the statistical parameter combination KRT-SKW using the SVM-DTC-LR method. In the 150-ml sample chamber experiment, the highest accuracy rate was obtained from the combination STD-AVG-KRT-SKW-MIN-MAX with the SVM-DTC-LR method. In the 250-ml sample chamber experiment, the highest accuracy rate was obtained from the combination STD-AVG-MIN-MAX with the SVM-DTC-ANN method. Otherwise, the lowest value was obtained by the combination of KRT-SKW and KRT-SKW-MIN-MAX in all the classification methods. These results occurred in the larger sample chambers 150-ml and 250-ml. This shows that the combination of KRT-SKW and KRT-SKW-MIN-MAX are affected by the amount of gas concentration. In this study, the statistical parameter combinations that were relatively unaffected by the sample chamber size were STD and AVG where even though they were combined with KRT-SKW or MIN-MAX parameters, the resulting accuracy was still quite good. It can be seen from the smallest result that the STD-AVG-KRT-SKW combination in the 250-ml sample chamber with the DTC method produces an accuracy value of 74.85%.

The results from the best combination of statistical parameters were then used as reference for the meta-learning method. The dark purple color in Table 7 shows the best series of statistical parameter combinations. The highest accuracy rate for the 50-ml sample chamber was 95.71%, for the 150-ml sample chamber it was 92.85%, and for the 250-ml sample chamber it was 87.14%. For more details, Fig. 9 displays a comparison graph of the best accuracy rates based on sample chamber and classifier method. From all comparisons, it can be seen that the accuracy rate of the 50-ml sample chamber (smallest size) was higher than that of the other chambers for all methods. The second best result was obtained with the 150-ml sample chamber (medium size), even though the accuracy was only 83.42% with the SVM method, below that

**Table 5**
Output Metrics of the SVM Method for 50-ml Sample Chamber.

| Class | Measurement | | | | |
|---|---|---|---|---|---|
| | Support | Precision | Recall | F1 Score | Accuracy |
| AS000 | 50 | 0.94 | 0.96 | 0.95 | 94.28% |
| AS010 | 50 | 0.89 | 0.96 | 0.92 | |
| AS025 | 50 | 0.98 | 0.96 | 0.97 | |
| AS050 | 50 | 0.98 | 0.92 | 0.95 | |
| AS075 | 50 | 0.92 | 0.92 | 0.92 | |
| AS090 | 50 | 0.92 | 0.94 | 0.93 | |
| AS100 | 50 | 0.98 | 0.94 | 0.96 | |

**Table 6**
Ensemble Learning Results for 50-ml Sample Chamber Using Voting Method.

| Method | SVM | DTC | LR | ANN | Ensemble |
|---|---|---|---|---|---|
| Accuracy | 94.57 | 91.14 | 93.71 | 92.85 | 95.71 |

**Table 7**
Comparison of Accuracy Between Classification Methods using a Combination of Statistical Parameters based on Sample Chamber Size.

| Sample Chamber | Statistical Parameter Combinations | Accuracy Result | | | | |
|---|---|---|---|---|---|---|
| | | SVM | DTC | LR | ANN | Ensemble |
| 50 ml | STD-AVG | 94.00 | 88.00 | 93.71 | 92.57 | |
| | STD-AVG-MIN-MAX | 94.28 | 88.85 | 93.71 | 93.41 | |
| | STD-AVG-KRT-SKW | 94.57 | 88.85 | 93.42 | 92.28 | |
| | STD-AVG-KRT-SKW-MIN-MAX | 94.28 | 89.71 | 93.14 | 92.28 | |
| | KRT-SKW | 94.57 | 91.14 | 93.71 | 92.85 | 95.71% |
| | KRT-SKW-MIN-MAX | 94.57 | 89.14 | 93.14 | 92.57 | |
| 150 ml | STD-AVG | 81.00 | 85.14 | 87.71 | 87.14 | |
| | STD-AVG-MIN-MAX | 80.57 | 83.42 | 87.14 | 86.00 | |
| | STD-AVG-KRT-SKW | 83.42 | 84.00 | 88.57 | 87.14 | |
| | STD-AVG-KRT-SKW-MIN-MAX | 83.42 | 85.42 | 88.57 | 86.85 | 92.85% |
| | KRT-SKW | 68.88 | 73.42 | 77.14 | 80.85 | |
| | KRT-SKW-MIN-MAX | 78.57 | 81.71 | 86.57 | 84.00 | |
| 250 ml | STD-AVG | 84.00 | 77.42 | 86.85 | 86.85 | |
| | STD-AVG-MIN-MAX | 84.28 | 79.71 | 86.57 | 86.85 | 87.14% |
| | STD-AVG-KRT-SKW | 80.85 | 74.85 | 83.71 | 83.14 | |
| | STD-AVG-KRT-SKW-MIN-MAX | 82.57 | 75.42 | 82.85 | 81.71 | |
| | KRT-SKW | 68.57 | 67.14 | 70.28 | 71.42 | |
| | KRT-SKW-MIN-MAX | 78.85 | 74.28 | 78.85 | 78.85 | |

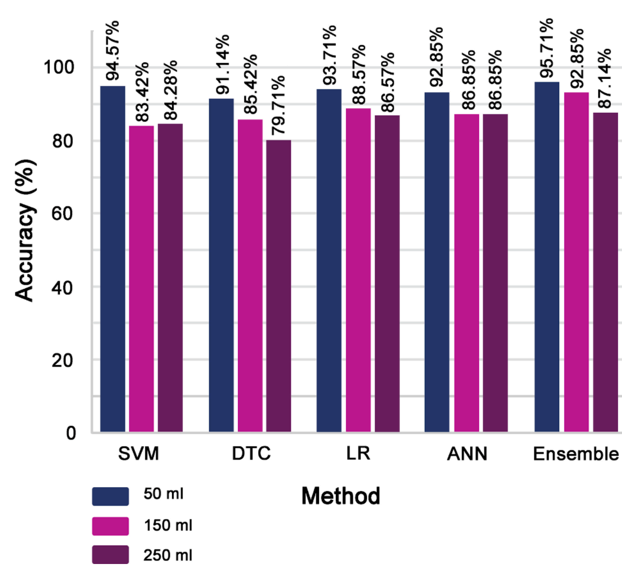


**Fig. 9.** Accuracy comparison chart between sample chamber and classifier method.

of the 250-ml sample chamber. Finally, the results from the 250-ml sample chamber (large size) were the lowest because the accuracy rates for the LR and DTC methods were lower than those of the other methods. These results indicate that the higher gas concentration in the smallest sample chamber positively influenced the classification results, where the gas concentration positively affected the quality of the data taken. This made the classification result more accurate.

## 5. Conclusions

In this study, three sample chambers of different sizes were used in an experiment to determine the effect of gas concentration on the detection and classification results. To get the optimum accuracy rate, a comparison between the results from different combinations of statistical parameters and 4 classification methods was used, which were further improved by using an ensemble method. The best classification results were obtained using the SVM method with KRT-SKW as statistical parameters, with an accuracy rate of 94.57%. After using the ensemble method, the best result was an accuracy of 95.71%. The classification comparison based on sample chamber size gave the following final accuracy rates: 95.71% for 50 ml, 92.85% for 150 ml, and 87.14% for 250 ml. It can be concluded that the higher the gas concentration, the better the classification results. In future work, the analysis will be extended by adding factors such as temperature, noise, and humidity.

## CRediT authorship contribution statement

**Sulaiman Wakhid:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing . **Shoffi Izza Sabilla:** Formal analysis, Data curation, Project administration, Resources, Software, Supervision, Writing – review & editing . **Riyanarto Sarno:** Conceptualization, Data curation, Funding acquisition, Investigation, Supervision, Validation, Writing – review & editing.

## 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

This research was funded by the Indonesian Ministry of Education and Culture under Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) Program, and Institut Teknologi Sepuluh Nopember (ITS) under project scheme of the Publication Writing and IPR Incentive Program (PPHKI).

## References

Ahmad, M.W., Mourshed, M., Rezgui, Y., 2017. Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy Build. 147, 77–89.

Ahmed, K.R., Akter, S., Marandi, A., Schüth, C., 2021. A simple and robust wetland classification approach by using optical indices, unsupervised and supervised machine learning algorithms. Remote Sens. Appl.: Soc. Environ. 23, 100569.

Alim, A., Song, H., Liu, Y., Zou, T., Zhang, Y., Zhang, S., Raza, A., 2019. Research of beef-meaty aroma compounds from yeast extract using carbon module labeling (CAMOLA) technique. LWT 112, 108239.

Cervantes, B., Monroy, R., Medina-Pérez, M.A., Gonzalez-Mendoza, M., Ramirez-Marquez, J., 2018. Some features speak loud, but together they all speak louder: A study on the correlation between classification error and feature usage in decision-tree classification ensembles. Eng. Appl. Artif. Intell. 67, 270–282.

Dent, B.B., Forbes, S.L., Stuart, B.H., 2004. Review of human decomposition processes in soil. Environ. Geol. 45 (4), 576–585.

Devi, S., Sahoo, M.N., Muhammad, K., Ding, W., Bakshi, S., 2019. Hiding medical information in brain MR images without affecting accuracy of classifying pathological brain. Futur. Gener. Comput. Syst. 99, 235–246.

Düntsch, I., Gediga, G., 2020. Indices for rough set approximation and the application to confusion matrices. Int. J. Approx. Reason. 118, 155–172.

Ghiasi, M.M., Zendehboudi, S., Mohsenipour, A.A., 2020. Decision tree-based diagnosis of coronary artery disease: CART model. Comput. Methods Programs Biomed. 192, 105400.

Ghosh, A., Dey, P., 2021. Flood Severity assessment of the coastal tract situated between Muriganga and Saptamukhi estuaries of Sundarban delta of India using Frequency Ratio (FR), Fuzzy Logic (FL), Logistic Regression (LR) and Random Forest (RF) models. Reg. Stud. Mar. Sci. 42, 101624.

Goudarzi, G., Hopke, P.K., Yazdani, M., 2021. Forecasting PM2.5 concentration using artificial neural network and its health effects in Ahvaz, Iran. Chemosphere 283, 131285.

Gu, S.Q., Wang, X.C., Tao, N.P., Wu, N., 2013. Characterization of volatile compounds in different edible parts of steamed Chinese mitten crab (Eriocheir sinensis). Food Res. Int. 54 (1), 81–92.

Han, L., Jinghao, J., Feixiang, Z., Guohua, H., 2015. Hairtail (Trichiurus haumela) freshness determination method based on electronic nose. J. Food Meas. Charact. 9 (4), 541–549.

Hariyanto, Sarno, R., Wijaya, D.R., 2018. Detection of diabetes from gas analysis of human breath using e-Nose. In: Proceedings of the 11th International Conference on Information and Communication Technology and System, ICTS 2017, vol. 2018-January, pp. 241–246.

Harsono, W., Sarno, R., Sabilla, S.I., 2020. Recognition of original arabica civet coffee based on odor using electronic nose and machine learning. In: Proceedings – 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, iSemantic 2020, vol. 2020, pp. 333–339.

Hasan, M.A., Sarno, R., Sabilla, S.I., 2020. Optimizing machine learning parameters for classifying the sweetness of pineapple aroma using electronic nose. Int. J. Intell. Eng. Syst. 13 (5), 122–132.

Hui, G., 2016. Apple analysis and the electronic nose. Electron. Noses Tongues Food Sci. 15–19.

Jiang, J., Li, J., Zheng, F., Lin, H., Hui, G., 2015. Rapid freshness analysis of mantis shrimps (Oratosquilla oratoria) by using electronic nose. J. Food Meas. Charact. 10 (1), 48–55.

Jiang, G., Wang, W., 2017. Error estimation based on variance analysis of k-fold cross-validation. Pattern Recogn. 69, 94–106.

Książek, W., Gandor, M., Pławiak, P., 2021. Comparison of various approaches to combine logistic regression with genetic algorithms in survival prediction of hepatocellular carcinoma. Comput. Biol. Med. 134, 104431.

Kumari, S., Kumar, D., Mittal, M., 2021. An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. Int. J. Cogn. Comput. Eng. 2, 40–46.

Laga, S.A., Sarno, R., 2019. Optimal sample temperature of electronic nose for detecting beef and pork mixture. In: 2019 International Conference on Information and Communications Technology, ICOIACT 2019, pp. 398–402.

Laga, S.A., Sarno, R., 2020. Temperature effect of electronic nose sampling for classifying mixture of beef and pork. Indones. J. Electr. Eng. Comput. Sci. 19 (3), 1626–1634.

Li, J., Feng, H., Liu, W., Gao, Y., Hui, G., Oct. 2016. Design of a portable electronic nose system and application in K Value prediction for large yellow croaker (Pseudosciaena crocea). Food Anal. Methods 9 (10), 2943–2951.

Li, J., Lin, M., 2021. Ensemble learning with diversified base models for fault diagnosis in nuclear power plants. Ann. Nucl. Energy 158, 108265.

Liu, Y., Zhang, F., Zhu, B., Ruan, X., Yi, X., Li, J., Gao, Y., Hui, G., 2020. Effect of sodium lactate coating enriched with nisin on beef strip loins (M. Longissimus lumborum) quality during cold storage and electronic nose rapid evaluation. J. Food Meas. Charact. 14 (6), 2998–3009.

Luque, A., Carrasco, A., Martín, A., De Las Heras, A., 2019. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. Pattern Recognit. 91, 216–231.

Malikhah, M., Sarno, R., Sabilla, S.I., 2021. Ensemble learning for optimizing classification of pork adulteration in beef based on electronic nose dataset. Int. J. Intell. Eng. Syst. 14 (4).

Mohammadi, M., Rashid, T.A., Karim, S.H.T., Aldalwie, A.H.M., Tho, Q.T., Bidaki, M., Rahmani, A.M., Hosseinzadeh, M., 2021. A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. J. Netw. Comput. Appl. 178, 102983.

Nakyinsige, K., Man, Y.B.C., Sazili, A.Q., 2012. Halal authenticity issues in meat and meat products. Meat Sci. 91 (3), 207–214.

Patel, H.K., 2014. "The Electronic Nose. Artificial Olfaction Technology".

Pedregosa, F., Michel, V., Grisel, O., et al., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Phillips, T., Abdulla, W., 2021. Developing a new ensemble approach with multi-class SVMs for Manuka honey quality classification. Appl. Soft Comput. 111, 107710.

Raza, K., 2019. Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule. U-Healthcare Monit. Syst. 179–196.

Rivai, M., Budiman, F., Purwanto, D., Simamora, J., 2018. Meat freshness identification system using gas sensor array and color sensor in conjunction with neural network pattern recognition. J. Theor. Appl. Inf. Technol. 30 (12).

Rustinsyah, R., 2019. The significance of social relations in rural development: A case study of a beef-cattle farmer group in Indonesia. J. Co-op. Organ. Manag. 7 (2), 100088.

Sabilla, S.I., Sarno, R., Triyana, K., Hayashi, K., 2020. Deep learning in a sensor array system based on the distribution of volatile compounds from meat cuts using GC–MS analysis. Sens. Bio-Sens. Res. 29, 100371.

Sarno, R., Triyana, K., Sabilla, S.I., Wijaya, D.R., Sunaryono, D., Fatichah, C., 2020. Detecting pork adulteration in beef for halal authentication using an optimized electronic nose system. IEEE Access.

Shao, C., Zheng, H., Zhou, Z., Li, J., Lou, X., Hui, G., Zhao, Z., 2018. Ridgetail white prawn (Exopalaemon carinicauda) K Value predicting method by using electronic nose combined with non-linear data analysis model. Food Anal. Methods 11 (11), 3121–3129.

Shen, H., Wei, T., Zhang, Z., Zheng, Q., Guo, R., Jiang, H., Zhang, G., Zheng, J., 2020. Discrimination of five brands of instant vermicelli seasonings by HS-SPME/GC-MS and electronic nose. J. Food Sci. Technol. 57 (11), 4160–4170.

Tian, X., Wang, J., Cui, S., 2013. Analysis of pork adulteration in minced mutton using electronic nose of metal oxide sensors. J. Food Eng. 119 (4), 744–749.

Wakhid, S., Sarno, R., Sabilla, S.I., Maghfira, D.B., 2020. Detection and classification of indonesian civet and non-civet coffee based on statistical analysis comparison using E-Nose. Int. J. Intell. Eng. Syst. 13 (4), 56–65.

Wang, H., Zhao, Z., Wang, Z., Xu, G., Wang, L., 2020. Independent component analysis-based baseline drift interference suppression of portable spectrometer for optical electronic nose of internet of things. IEEE Trans. Ind. Inform. 16 (4), 2698–2706.

Wijaya, D.R., Sarno, R., Daiva, A.F., 2017. Electronic nose for classifying beef and pork using Naïve Bayes. In: Proceedings – 2017 International Seminar on Sensor, Instrumentation, Measurement and Metrology: Innovation for the Advancement and Competitiveness of the Nation, ISSIMM 2017, 2017, vol. 2017-January, pp. 104–108.

Wijaya, D.R., Sarno, R., Zulaika, E., Sabila, S.I., 2017. Development of mobile electronic nose for beef quality monitoring. Procedia Comput. Sci. 124, 728–735.

Wijaya, D.R., Sarno, R., Zulaika, E., 2019. Noise filtering framework for electronic nose signals: an application for beef quality monitoring. Comput. Electron. Agric. 157, 305–321.

Wijaya, D.R., Sarno, R., Zulaika, E., 2021. DWTLSTM for electronic nose signal processing in beef quality monitoring. Sensors Actuators B Chem. 326, 128931.

Wong, T.T., 2017. Parametric methods for comparing the performance of two classification algorithms evaluated by k-fold cross validation on multiple data sets. Pattern Recogn. 65, 97–107.

Zheng, H., Wang, S., Ping, X., et al., 2019. Study of spinyhead croaker (Collichthys lucidus) fat content forecasting model based on electronic nose and non-linear data resolution model. Food Anal. Methods 12 (9), 1927–1937.

Zheng, H., Ying, X., Wang, W., et al., 2019. Study of sensitivity evaluation on ridgetail white prawn (Exopalaemon carinicauda) quality examination methods. Int. J. Food Prop. 22 (1), 942–951.

Zhiyi, H., Chenchao, H., Jiajia, Z., Jian, L., Guohua, H., 2017. Electronic nose system fabrication and application in large yellow croaker (Pseudosciaena crocea) fressness prediction. J. Food Meas. Charact. 11 (1), 33–40.

Zhou, M., Khir, R., Pan, Z., Campbell, J.F., Mutters, R., Hu, Z., May 2021. Feasibility of detection of infested rice using an electronic nose. J. Stored Prod. Res. 92, 101805.