E-nose: a low-cost fruit ripeness monitoring system

Pankaj Tyagi,[1] Rahul Semwal,[1] Anju Sharma,[2] Uma Shanker Tiwary,[1] Pritish Varadwaj[2]

[1]Department of Information Technology, Indian Institute of Information Technology Allahabad, Prayagraj;
[2]Department of Applied Sciences, Indian Institute of Information Technology Allahabad, Prayagraj, India

Abstract

All fruits emit some specific volatile organic compounds (VOCs) during their life cycle. These VOCs have specific characteristics; by using these characteristics fruit ripening stage can be identified without destroying the fruit. In this study, an application-specific electronic nose device was designed for monitoring fruit ripeness.

The proposed electronic nose is cost-efficient and does not require any modern or costly laboratory instruments. Metal oxide semiconductor (MOS) sensors were used for designing the proposed electronic nose. These MOS sensors were integrated with a microcontroller board to detect and extract the meaningful features of VOCs, and an artificial neural network (ANN) algorithm was used for pattern recognition. Measurements were done with apples, bananas, oranges, grapes, and pomegranates. The designed electronic nose proved reliable in classifying fruit samples into three different fruit ripening stages (unripe, ripe, and over-ripe) with high precision and recall. Furthermore, the proposed electronic nose performed uniformly on all three fruit ripening stages with an average accuracy of ≥95%.

Correspondence: Pritish Varadwaj, Department of Applied Sciences, Indian Institute of Information Technology Allahabad, Prayagraj, Uttarpradesh, 211012, India.
Tel.: +91.532.2922090.
E-mail: pritish@iiita.ac.in

Key words: artificial neural network; electronic nose; fruit quality control; MOS sensors; volatile organic compounds.

Acknowledgments: the electronic-nose was designed by the hardware components provided through the Design Innovation Centre (SPOKE) - IIIT-Allahabad. The authors would also like to express their sincere appreciation toward Dr. Suneel Yadav, DIC spoke coordinator IIIT-Allahabad, for his valuable support during our E-nose development.

Contributions: the main ideas, the design of the electronic-nose and code implementation were delivered by PT. RS and AS helped in Data collection and manuscript preparation while PV and UST supervised the entire research work and edited the manuscript.

Conflict of interest: the authors declare no conflict of interests.

Funding: none.

Availability of data and materials: fruits odor data and codes are available on the institute server https://bioserver.iiita.ac.in/

Received for publication: 23 February 2022.
Revision received: 19 May 2022.
Accepted for publication: 10 July 2022.

Journal of Agricultural Engineering 2023; LIV:1389
doi:10.4081/jae.2022.1389

Introduction

Monitoring and controlling fruit ripeness is an essential task in fruit farming. The quality of fruit in the consumer market largely depends on the fruit's maturity or ripeness during harvesting, storage, and market distribution. In the past, several techniques have been proposed for fruit ripeness monitoring, however, these techniques come with major disadvantages; i) the fruit samples are often destroyed while analysing ripeness status, ii) some of these techniques are not practical for farming and storage. Because of these drawbacks, most techniques are unsuitable for fruit ripeness monitoring; therefore, human involvement is required to obtain fruit storage life and optimal harvest dates. Due to subjective interpretation, an enormous quantity of fruits are reaped too early or too late, reaching the markets in bad condition (Brezmes *et al*., 2005).

Presently, there are two commonly used destructive methods for fruit ripeness monitoring; i) the sensory-evaluation method and ii) the physicochemical-indexes detection method (Singh and Singh, 1994; Xu *et al*., 2016). The first method requires human observation to evaluate the fruit ripeness using human vision, taste, and smell. This method relies on human perception; therefore, it suffers from low evaluation speed and is highly influenced by subjective factors (Zhaoqi *et al*., 2002). The second method evaluates the fruit ripeness by extracting the fruit's physical and chemical features. It is an effective fruit ripeness detection method because it does not require human intervention, but the operating procedure of this method makes it non-ideal for fruit ripeness monitoring for the massive fruit production industry. In this method, sample fruits need to be damaged to extract the physical and chemical features, and this method also suffers from low evaluation speed (Singh *et al*., 2010).

Furthermore, there are two non-destructive methods for fruit ripeness monitoring: the near-infrared spectroscopy method (Khodabakhshian *et al*., 2017) and the electronic nose detection method (Chen *et al*., 2018; Voss *et al*., 2019; Tan and Xu, 2020). The near-infrared spectroscopy method extracts the internal and external features of a fruit sample. However, in this method, fea-

ture extraction is very difficult during storage due to mutual shadowing among fruits, which makes this method inefficient for fruit ripeness monitoring during storage. The biological olfaction process influences the electronic nose fruit ripeness monitoring method (Beghi *et al*., 2017). It detects and analyses the volatile organic compounds (VOC) emitted by the fruits during different developmental phases, which makes this method more suitable for fruit ripeness monitoring without being influenced by field angle and subjective factors. Presently, the E-nose technology has been successfully used in many research areas, such as quality monitoring of vegetables (Giovenzana *et al*., 2014; Yang *et al*., 2020), quality control of nuts (Yoshida *et al*., 2012), detection of meat spoilage (Kodogiannis, 2017), and quality monitoring of milk (Poghossian *et al*., 2019). Applying the E-nose technology can offer a non-destructive practical method for real-time fruit ripeness monitoring in mass production.

The electronic nose is a device that simulates the biological sense of smell (Baietto and Wilson, 2015). The E-nose device is developed to detect and differentiate among complex odorants using an array of metal oxide semiconductor sensors. The sensor array generates a distinctive electronic fingerprint under exposure to an odorant stimulus (Ghasemi *et al*., 2011). The responses of all the sensors in electronic fingerprints are recorded, and various machine-learning algorithms can be applied to these fingerprints to discriminate among different fruit ripeness stages. These characteristics of the E-nose substantially suggest using an electronic nose to provide a quick perception of the VOC (Tang *et al*., 2010). A typical electronic nose mainly consists of three parts (Pearce, 1997; Haugen and Kvaal, 1998; Yu *et al*., 2008): i) sample delivery unit; fruit samples are placed in this unit for a particular time to collect VOC emitted by the fruit samples; ii) sensor unit; multiple sensors are placed in this unit to extract the information from the VOC collected in the sample delivery system; iii) data acquisition and computational unit; extracted information from the sensor unit is analysed by the computational unit to generate a smell fingerprint.

The motive of this study was: i) to design and integrate a low-cost electronic nose device using a sensor array of five metal oxide semiconductor (MOS) sensors and one digital temperature and humidity (DHT) sensor to classify fruits into three distinct ripeness *unripe*, *ripe*, and *over-ripe* stages. Several research studies have described the application of an electronic nose for fruit quality identification and ripeness monitoring (Mavani *et al*., 2021); though all these studies target a specific fruit, in our study, we have performed our experiment on multiple types of fruits; ii) to evaluate the performance of the developed electronic nose device coupled with an artificial neural network.

Materials and Methods

Electronic nose design

In this paper, a portable electronic nose was designed using five MOS sensors and one DHT sensor. All the sensors were integrated with a microcontroller board (Arduino Uno). The open-source Arduino Software (IDE) was used to write the program code, and later these codes were uploaded to the microcontroller board. Finally, the program interacted with the microcontroller board to control the sampling process, read and interpret the sensor responses, and store these responses in a file for later analysis.

The E-nose designed in this study comprises three main modules: sample delivery unit, sensor unit, and data-acquisition and computational unit. Details for each of these components are described in the following sections. The basic structure of an electronic nose is shown in Figure 1. The architecture of the designed electronic nose in this study is shown in Figure 2.

Sample delivery unit

The sample delivery unit was a 1 cubic ft cylindrical iron chamber equipped with one outlet valve at the top of the chamber, connected with a 4.6 CFM REX 30 vacuum pump which was used to remove the fruit aroma of the previous sample from the sampling unit. One inlet valve was also present at the top of this chamber to supply the fresh air inside the chamber. This chamber was also equipped with a USB cable with the help of a rubber stopper; one end of the USB cable was placed inside the chamber, and the other end was placed outside the chamber to connect the sensor unit to the computer.

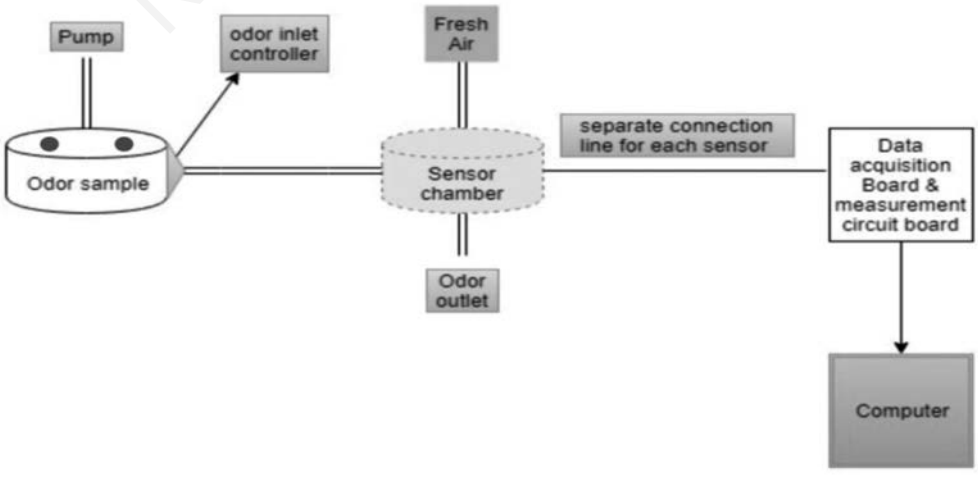

Figure 1. The basic architecture of an electronic-nose.

Sensor unit

Five MOS sensors and one DHT sensor were used in this study. All the sensors were symmetrically integrated on a breadboard and connected to a microcontroller board. In this study, sensors were selected based on their sensitivity to ensure the selectivity of the designed electronic nose. However, there are no specific rules to determine the numbers or sensor selection criteria (Schaller *et al*., 1998). The sensor unit was placed inside the sample delivery chamber and connected to a computer via a USB cable. Table 1 lists the various sensors used in this study to design an E-nose and describes their applications and detection limits.

MOS sensors function accurately at a specific temperature (40°C). To acquire this temperature, MOS sensors have a heater resistance built inside the sensor. The heater and measurement circuits need to be connected with a constant DC voltage across the heater and measurement pins for a specific time (≈48 Hr) before data collection. The sensors get sensitive to a specific odor present in the air at this raised temperature. Sensor conductivity changes in the presence of this detectable odour, depending on the odour and sensor type (Fine *et al*., 2010). This change in the sensor conductance is measured by the measuring circuit in the form of a voltage signal. Change in the sensor conductance in various conditions is shown in Table 2. Figure 3 describes the measuring circuits for the TGS26xx and MQ-x sensors, respectively. The TGS 26xx, MQ-4, and MQ-5 sensors require one input voltage (5v) for both the

Figure 2. The designed electronic nose.

Table 1. Sensors used in this study to design the E-nose device.

Number	Name	Target odors	Sensitivity
S1	MQ-5	Natural gas, alcohol, and hydrogen	200–10000 ppm
S2	MQ-4	Methane	300–10000 ppm
S3	TGS 2620	Alcohol and solvent vapors	50–5000 ppm
S4	TGS 2610	LPG and component gases	50–5000 ppm
S5	TGS 2602	Air contaminants (VOCs and odorous gases)	1–30 ppm
S6	DHT 11	Humidity and temperature sensor	

LPG, liquefied petroleum gas; VOCs, volatile organic compounds.

Table 2. Change in sensor conductance in different conditions.

Sensor type	Oxidizing gases	Reducing gases
n-type	Conductivity decrease	Conductivity increase
p-type	Conductivity increase	Conductivity decrease

heater circuit (V_H) and the measuring circuit (V_C). Sensor response of TGS 26xx, MQ-4, and MQ-5 can be measured by applying 1 kΩ, 10 kΩ, and 20 kΩ load resistance (R_L), respectively, across V_{out} pin (Mamat *et al*., 2011).

Data-acquisition unit

The data-acquisition unit comprises a microcontroller board (Arduino Uno), an open-source Arduino Uno Software (IDE), and a computer. First, the microcontroller board was connected to the computer through a USB cable. Then, a program was written on Arduino Uno Software (IDE) and uploaded to the microcontroller board to read sensor voltage and to receive and transmit data to a file.

Data collection

Apples, bananas, oranges, grapes, and pomegranates fruits were used in this study for fruit ripeness monitoring. These fruits were chosen for this study based on the availability of unripe fruits in the local market. To define the fruit ripeness stages, we relied on human experience, and human observations, vision and smell were used to define the ripe, unripe, and overripe fruits. Unripe bananas, grapes, and apples are easy to find in the Indian consumer market because these fruits are also used as vegetables and pickles.

However, finding unripe oranges and pomegranates was a time-consuming task, and for that, we visited the consumer markets on multiple days and selected a few, sometimes limited to one unripe fruit on a single day.

All the fruits, in two different categories, unripe and ripe, were purchased from the local market on various days. For unripe and ripe fruits, data collection was performed on the same day the fruit was purchased. For overripe fruits, data collection was performed when the available ripe fruits reached the overripe stage. For each fruit, 30 samples (10 unripe, 10 ripe, and 10 overripe) were used. In this study, experiments were carried out on five fruits (apple, banana, orange, grape, and pomegranate), thus a total of 150 fruit samples were used. For each fruit sample, 120 data samples were recorded, thus 6000 data samples were collected in each category (unripe, ripe, and overripe). To ensure the specificity of our machine learning model, we included 6000 sample recordings of non-fruit odor (negative data) in our dataset, thus our final dataset contained 24,000 (3 category*6,000 sample+6,000 non-fruit odor sample) data samples.

The sensor array response curve to different states of fruit ripeness and the sensor array response curve to multiple fruit types in the same ripeness state are shown in Figure 4. In this study, we considered the ripening state of fruits as an aspect, instead of the fruit type. Every fruit emits some VOCs during its life cycle, some VOCs (aldehydes, alcohol, ketone, *etc*.) are common to all fruits, and some VOCs are specific to a fruit (El Hadi *et al*., 2013). Our objective was to capture these common volatile organic compounds to classify multiple fruits into their ripeness states. To achieve this objective, multiple combinations of sensors were tested, and the sensor array listed in this paper was able to capture the common VOCs among the fruits. The sensor array response curve for ripe apple, ripe banana, ripe grape, ripe orange, and ripe pomegranate are shown in Figure 4B and E-H, respectively, and the curve patterns of these graphs are very similar to each other which indicate the ability of the designed E-nose to capture the ripeness state of these fruits without considering the type of fruit as an aspect.

Measurement protocol

The sensor unit was connected to a computer via a USB cable, and all the sensors were heated in ambient air for 48 hours before data collection to acquire a 40°C sensor temperature, thus producing sensitivity to a particular odour. Measurements were done for each fruit sample by following these two steps.

Step 1: Cleaning phase – the measurement begins by conducting the baseline voltage correction for sensors. In this process, air containing odour from the previous measurement was sucked out for 360 seconds, followed by the release of fresh ambient air into the chamber for 60 seceonds, thus obtaining a steady baseline voltage. The sensor voltage was recorded and saved in a file for later use in data preprocessing.

Step 2: Sampling phase – to accumulate adequate odour concentration comparable with the sensor sensitivity, the fruit sample was placed inside the chamber for 40 minutes, and then sample recording was done for 200 seconds. Data sample recording was repeated every 5 seconds; thus, each sampling phase contains 40

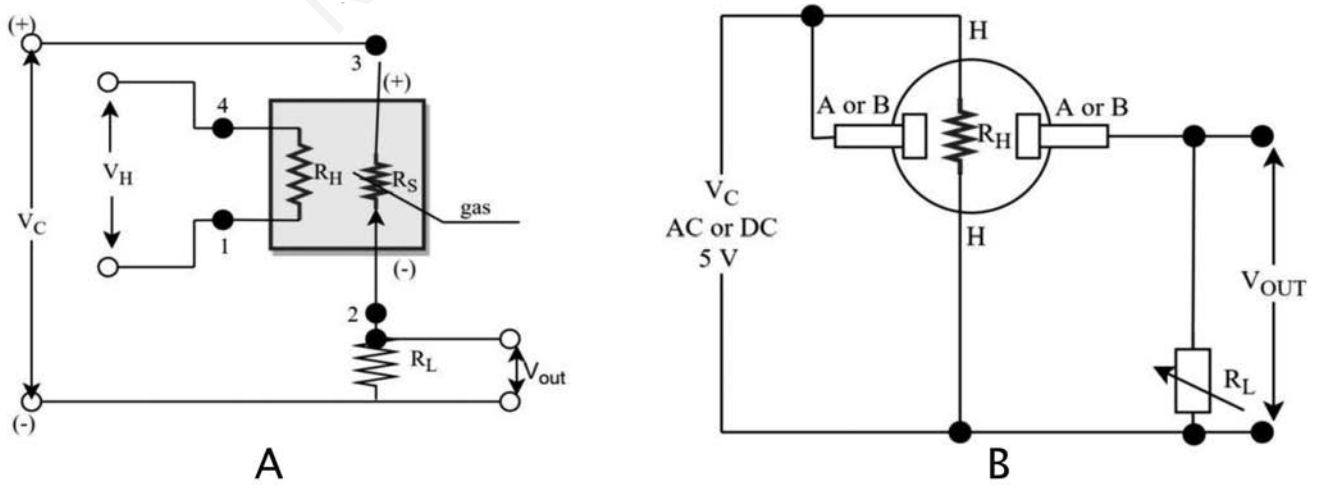

Figure 3. The measurement circuit for Metal oxide semiconductor sensors. A) TGS 26xx sensors measurement circuit; B) MQ –x sensors measurement circuit.

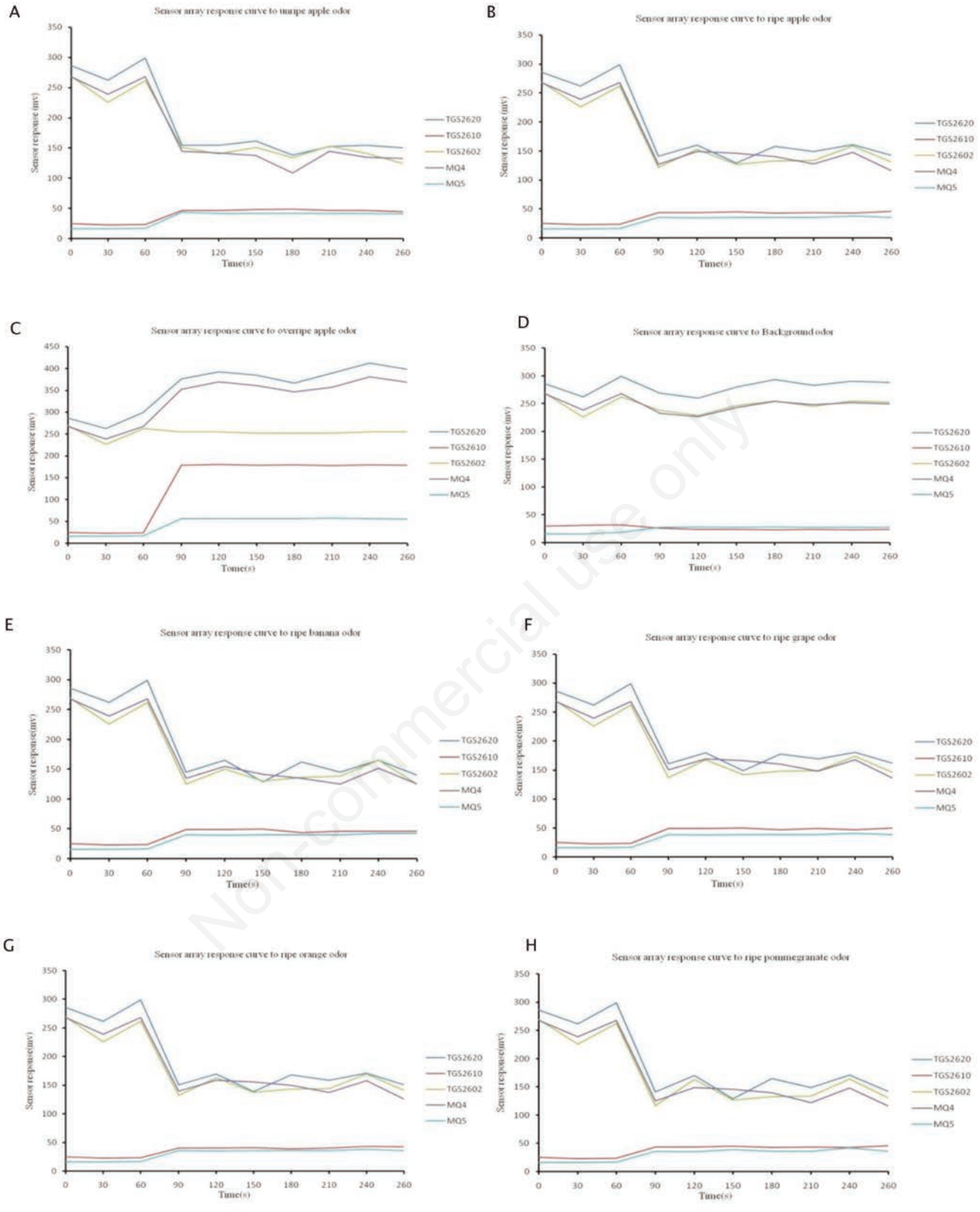

Figure 4. The sensor array response curve to different states of fruit ripeness. In sensor array response curve, sensor responses for first 60 seconds represents the baseline voltage corresponding to each sensor, and sensor responses for remaining 200 seconds are the sensor responses in presence of fruit odors. A) Sensor response to unripe apple odors; B) Sensor response to ripe apple odors; C) Sensor response to overripe apple odors; D) Sensor response to background odors; E) Sensor response to ripe banana odors; F) Sensor response to ripe grape odors; G) Sensor response to ripe orange odors; H) Sensor response to ripe pomegranate odors.

data samples. For each fruit sample, the cleaning and sampling process was repeated 3 times; therefore, 120 data samples were collected for each fruit sample. A total of 150 fruit samples were used in this study; thus, the process defined above was repeated 450 times, and as a consequence, the fruit odour data comprise 18000 samples. The cleaning and sampling time for the measurement process was obtained from sensor responses that were tested before the data collection. It was observed that 420 seconds for the cleaning phase and 40 minutes for odor accumulation was adequate. To collect non-fruit odor data, the experiment was performed over a long period of time in the open air without implementing the cleaning process. During the experiment, the temperature and humidity of the chamber were also recorded.

Data preprocessing

The prediction ability of machine learning algorithms largely depends on the data quality to obtain a generalised prediction model of the classification problem (Singh and Singh, 2019). Many research studies have presented the significance of data preprocessing to enhance the data quality and subsequently the classification performance of the machine learning algorithm. In these studies, several techniques have been suggested for data preprocessing in MOS sensors-based electronic nose (Sanaeifar *et al*., 2014). In our study, we applied the fractional normalisation method to our dataset as shown in Equation (1). In this method, the response for each sensor in the presence of fruit odor was subtracted and then divided from the sensors' baseline voltage reading recorded before each batch of sampling, hence obtaining relative responses of sensors. These relative responses of sensors were used to train and test our machine-learning model.

$$Vs_i = \frac{Vo_i - Vb_i}{Vb_i} \quad (1)$$

where the relative response of the ith sensor is V_{Si}, V_{oi} represents the ith sensor's response in the presence of fruit odour, and Vbi represents the baseline voltage of the ith sensor.

Pattern recognition

This study used an artificial neural network algorithm to obtain a generalised prediction model. The developed E-nose comprised 5 MOS sensors and hence collected 5 features for each fruit sample at a time. These feature sets were applied to train our artificial neural networks (ANN) classification model to predict the ripeness of fruit. An ANN model incorporates many layers of several processing units known as artificial neurons. A basic ANN model has three layers of artificial neurons: 1 output layer, 1 hidden layer, and 1 input layer. The number of hidden layers in a network can be increased according to the requirement. Figure 5 describes the architecture of this study's designed ANN model, which performed better than all the other models. The number of neurons in hidden layers and the number of hidden layers were obtained by trial and error. The ANN model was designed in Python (version 3.8.0) programming language using Keras (version 2.3.0) neural-network library, which was implemented using TensorFlow (version 2.0.0) machine learning platform.

The designed artificial neural network model contains 1 output layer, two hidden layers, and 1 input layer. Sensor array responses were presented to the input layer of the network; hence the input layer comprised 5 input neurons. The first hidden layer consisted of 10 neurons, while the second hidden layer included 8 neurons.

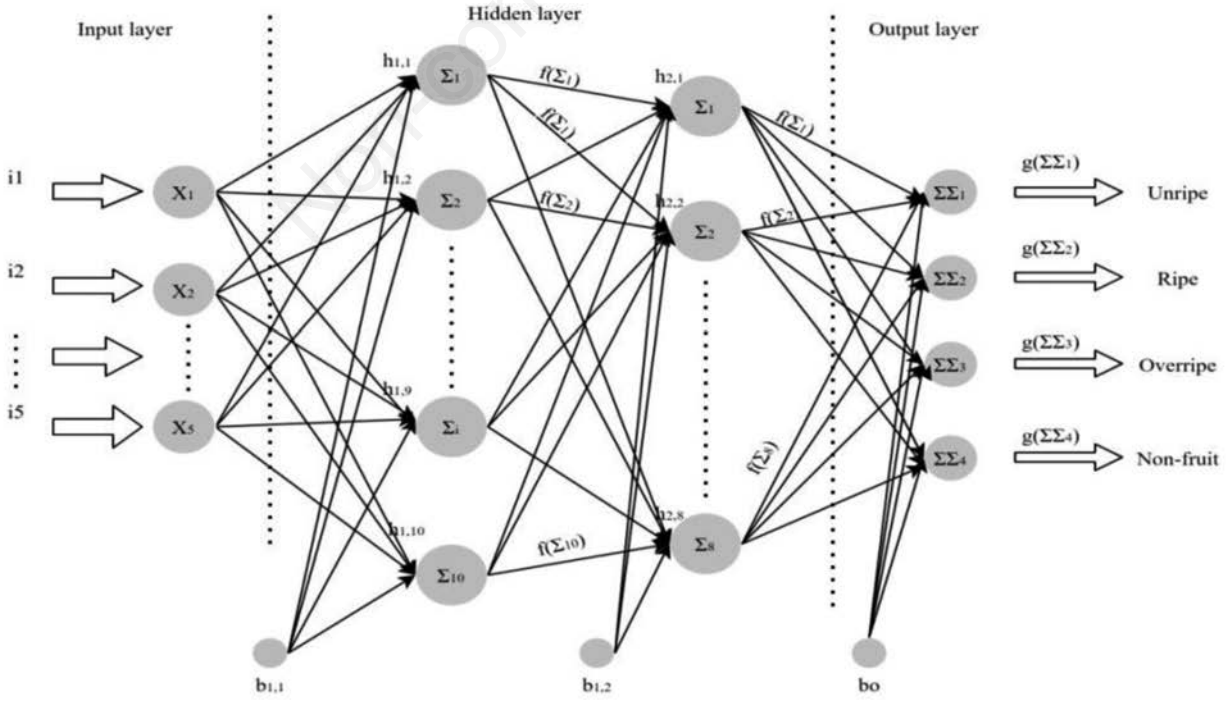

Figure 5. Structure of the designed artificial neural networks model.

The output layer included 4 output neurons; each neuron representing a distinct class. The ANN model parameters used in this study to train our feed-forward ANN model using the back-propagation algorithm are shown in Table 3.

The presentation of the entire training dataset to the feed-forward ANN model is known as one epoch while the presence of one set of features to the model is termed as iteration.

Equation (2) represents the k[th] neuron's actual value in the neural network.

$$N_k^a = \sum_{i=1}^{n} x_i^j w_{ik} + b_j \quad (2)$$

where w_{ik} is the weight between neuron i and neuron k. b_j represents the weighed sum of the bias node, n represents the number of inputs presented to the i[th] neuron. This study used the ReLU and Softmax activation functions for hidden and output layers, respectively. The output value of neurons in hidden layers and output layers are given in equation (3) and equation (4), respectively.

$$O_k^a = \max(0, N_k^a) \quad (3)$$

$$P(Y = j \mid \theta^i) = \frac{e^{\theta(i)}}{\sum_{j=0}^{k} e^{\theta_k^{(i)}}} \quad (4)$$

where $\Theta = W_0X_0 + W_1X_1 + W_2X2 + \ldots\ldots + W_kX_k + b_j = W^TX$. P is the probability of j[th] class over all possible target classes. O_m is the network's predicted output, which can be obtained by calculating the net value of the output layer and the net values of all the hidden layers. Equation (5) represents the error in the network, which is calculated by subtracting the obtained output (O_m) from the expected output (O_e) for that iteration. Total error and mean squared error (mse) are defined by equation (6) and equation (7), respectively.

$$E_m = O_e - O_m \quad (5)$$

$$E_{total} = \frac{1}{2}\sum_{m=1} E_m^2 \quad (6)$$

$$mse = \frac{1}{n}\sum_{k=1}^{n} E_{total} \quad (7)$$

where E_m is the error in the network for the i[th] iteration. The back-propagation algorithm using a gradient descent approach, aims to minimise the error given by Eq. (5). The back-propagation algorithm updates the network weights based on the error rate obtained in the previous epoch. First, it calculates the weight changes, and then network weights and thresholds are updated [Equation (8)-(10)] (Adak and Yumusak, 2016).

$$\delta_k = O_k(1 - O_k)E_m \quad (8)$$

$$\Delta w_{jk}^n = l_r \delta_k x_k + \mu \Delta w_{jk}^{n-1} \quad (9)$$

$$w_{jk}^n = w_{jk}^{n-1} + \Delta w_{jk}^n \quad (10)$$

where δ_k is the error rate for the k[th] node in the output layer of the network, O_k represents the output for k[th] neuron, Δw_{jk} represents the weight optimization rate between the j[th] and k[th] neurons. To prevent the network from settling in a local minimum, a momentum (μ) hyperparameter was used. While the learning rate (*lr*) decides how quickly the network learns in each iteration. In this study, the total number of epoch was set to 1,000, and the dataset collected from the fruit odour was presented to the ANN model.

Performance evaluation parameters

Various statistical and graphical measures were obtained to evaluate the performance of our classification model. To obtain these statistical scores, the predicted output of every test set needs to be labelled as true-positive, true-negative, false-positive, and false-negative. Each test sample is placed into one of these categories in a binary classification problem based on their actual class level and predicted class level. To obtain these levels for all classes, a multi-class classifier applies the one *vs.* all (OVA) or one *vs.* one (OVO) approach (Yang *et al*., 2013). In this study, the OVA approach was implemented. The OVA strategy splits a multi-class classification problem into one binary classification problem per class. Following the defined approach, true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) predictions were calculated for all the individual classes. In this study, the following statistical measures were obtained to assess the classification ability of the ANN model.

Precision – Precision is the ratio between correctly predicted positive observations and the total predicted positive observations, as shown in equation (11).

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

Recall is the ratio between correctly predicted positive observations and the total predictions in the respective class, as shown in equation (12).

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

F1-score – f1-score is calculated as the harmonic mean of precision and recall, as shown in equation (13).

$$F1 - score = \frac{2 * (Precision * Recall)}{(Recall + Precision)} \quad (13)$$

Table 3. Artificial neural networks parameters used in this study.

Parameters	Value
The hidden layers activation function	ReLU
Output layer activation function	Softmax
Learning rate	0.000001
Decay	1e-6
Momentum	0.9
Optimizer	SGD
Batch size	10
Validation split	0.20
Stopping rule (epoch)	1000

Accuracy – the ratio between correct predictions and the total number of test sets is known as accuracy, as described in equation (14).

$$Accuracy = \frac{TP + TN}{Total\ Samples} \quad (14)$$

Results and Discussion

A fruit aroma dataset comprising 24,000 sample sets was collected from the designed electronic nose,

following which a pre-processing technique was applied to this entire dataset as previously described. The obtained dataset was divided into two parts. The ANN model was trained using the first 80% dataset and the remaining 20% was used as the test dataset. The ANN model using the back-propagation algorithm was trained for 1000 epochs. Training and test cases were replicated 20 times. The training accuracy and loss, and test accuracy and loss were obtained at the end of the 20th run, as shown in Figure 6. The performance of the designed ANN model to predict fruit ripeness stages is displayed in Table 4. The designed ANN model was able to classify all samples into their respective classes with an average accuracy of 95%.

However, according to Wijaya *et al*. (Wijaya *et al*., 2017), accuracy alone is not a good performance evaluation measure for a machine learning model; therefore, precision, recall, and f1-score were also obtained for all four corresponding classes. Table 5 illustrates the statistical scores corresponding to all classes for the designed ANN model.

Table 5 shows that our model was able to differentiate non-

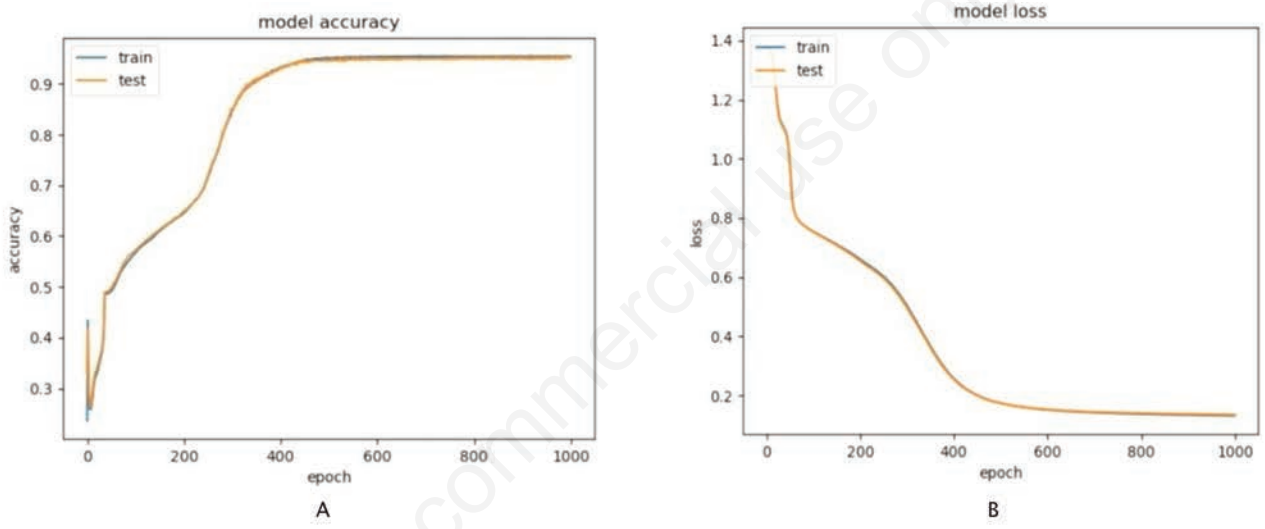

Figure 6. The performance curve of the designed artificial neural networks (ANN) model. A) Training and test accuracy curve of the designed ANN model; B) Training and test loss curve of the designed ANN model.

Table 4. Artificial neural networks classification results.

Fruit type	Test samples	Predicted as Class 0 (unripe)	Predicted as Class 1 (ripe)	Predicted as Class 2 (overripe)	Predicted as Class 3 (non-fruit)	Success rate (%)
Unripe	1335	1205	130	0	0	90.26
Ripe	1296	105	1191	0	0	91.89
Overripe	1314	8	0	1306	0	99.39
Non-fruit	1249	0	1	0	1248	99.91

Table 5. Performance evaluation based on statistical scores.

	Class 0 (unripe)	Class 1 (ripe)	Class 2 (overripe)	Class 3 (non-fruit)
Accuracy (%)	90.26	91.89	99.39	99.91
Precision (%)	91	91	100	100
Recall (%)	90	92	99	100
F1-score (%)	91	91	100	100

fruit odors from fruit odors with high accuracy (>99%). Similarly, it was able to differentiate overripe fruits from non-overripe fruits with high accuracy (>99%). At the same time, our model was able to classify ripe and unripe fruits into their respective categories with slightly low accuracy (>90%). The precision-recall and receiver operating characteristic (ROC) curves were also obtained to assess the classification ability of the ANN model.

Receiver operating characteristics analysis

Receiver operating characteristics (ROC) curve depicts the relationship between true positive rate (sensitivity) and false-positive rate (1-specificity), where true positive rate and false positive rate were plotted on the y-axis and x-axis, at various thresholds, respectively (Wang *et al*., 2006). The true positive rate defines the true-positive outcomes out of total positive samples while the false positive rate defines the false-positive outcomes out of total negative samples (Karami *et al*., 2020). In ROC analysis, the area under the curve (AUC) represents the test accuracy, *i.e.*, the accuracy is directly proportional to the AUC. The diagonal line [coordinates (0, 0) to (1, 1)] represents the random classification. Hence, the ROC curve should be above the diagonal line to accept a generalised classification model. In this study, the ROC curve corresponding to all four classes was obtained using the sklearn 'roc_curve, auc' package (Pedregosa *et al*., 2011). Figure 7 depicts the ROC curve with AUC (>0.98) corresponding to all four classes. In Figure 7, the ROC curve of the overripe class and non-fruit class were overlapping with each other; therefore, the overlapping section was zoomed and extracted (Figure 7b).

Precision-recall curve analysis

The precision-recall curve depicts the relationship between precision and recall, where precision and recall are plotted on

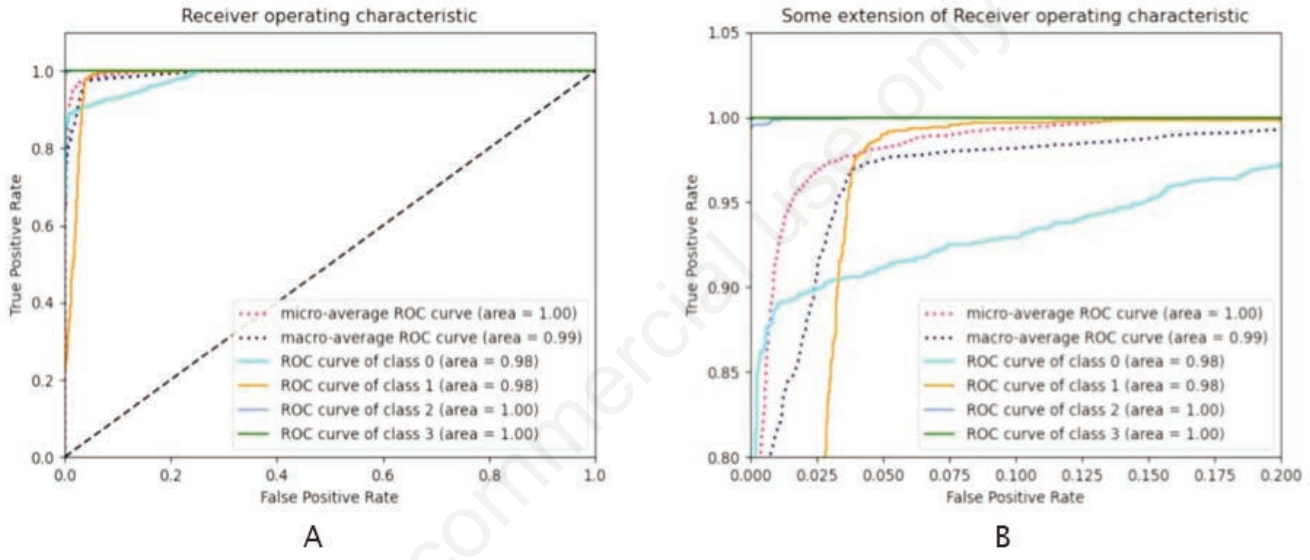

Figure 7. Receiver operating characteristics (ROC) curve analysis of the designed artificial neural networks model in this study: A) The ROC curve; B) Zoomed view of the ROC curve.

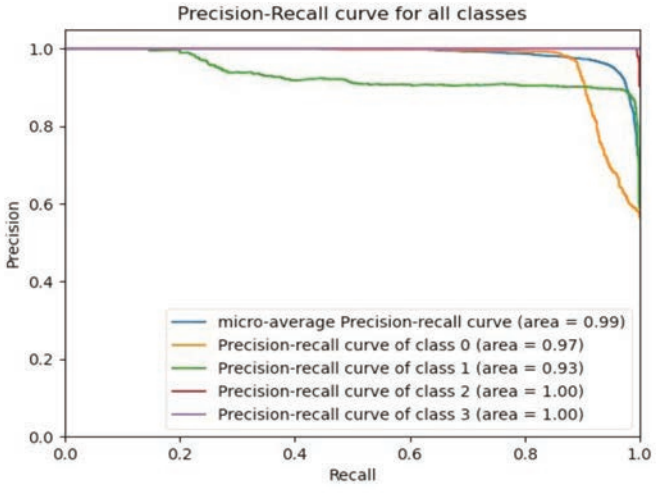

Figure 8. The precision-recall curve analysis of the designed artificial neural networks model.

Y-axis and X-axis at various thresholds, respectively. The greater area under the curve constitutes a high precision and high recall, where high recall indicates a low false-negative rate, and high precision indicates a low false-positive rate (Semwal *et al*., 2021). Therefore, the area under the precision-recall curve should be close to 1 square unit to obtain a good classifier. In this study, the precision-recall curve was obtained using the sklearn 'precision_recall_curve, auc' package (Pedregosa *et al*., 2011) with AUC (>0.93) corresponding to all four classes as depicted in Figure 8. Further, the micro average precision-recall curve was also plotted to demonstrate the average behaviour of the classifier (Figure 8).

Conclusions

This study presents an electronic nose system designed using five MOS sensors. The designed E-nose system was able to monitor the changes in odour fingerprints during the ripening of bananas, apples, grapes, oranges, and pomegranates. To ensure specificity, non-fruit odour data were also recorded.

The ANN machine learning model was used to discriminate fruits into three different ripening stages. Several statistical and graphical measures were taken into consideration to evaluate the classification model's performance. Results suggest that the designed E-nose could classify these dedicated fruits into their respective ripeness category with a high f1-score and an average accuracy of 95%. Furthermore, the in-house designed low-cost E-nose device has the potential to be used in the fruit storage facility to monitor the fruit ripeness and remove the overripe fruit before they damage other fruits.

The overall development cost of the E-nose device, excluding the sample delivery unit, was only USD 65. The developed Electronic-nose device is a prototype that has been tested in the laboratory in a closed environment hence, it can be categorised into technology readiness level 4. In the future, research can be done to choose a much more sophisticated sensor array to monitor fruit ripeness of all other fruits.

References

Adak M.F., Yumusak N. 2016. Classification of E-nose aroma data of four fruit types by ABC-based neural network. Sensors. 16:304.

Baietto M., Wilson A.D. 2015. Electronic-nose applications for fruit identification, ripeness and quality grading. Sensors. 15:899-931.

Beghi R., Buratti S., Giovenzana V., Benedetti S., Guidetti R. 2017. Electronic nose and visible-near infrared spectroscopy in fruit and vegetable monitoring. Rev. Anal. Chem. 36:4.

Brezmes J., Fructuoso M.L., Llobet E., Vilanova X., Recasens I., Orts J., Saiz G., Correig X. 2005. Evaluation of an electronic nose to assess fruit ripeness. IEEE. Sens. J. 5:97-108.

Chen L.Y., Wu C., Chou T., Chiu S.W., Tang K.T. 2018. Development of a dual MOS electronic nose/camera system for improving fruit ripeness classification. Sensors. 18:3256.

El Hadi M.A.M., Zhang F.J., Wu F.F., Zhou C.H., Tao J. 2013. Advances in fruit aroma volatile research. Molecules. 18:8200-29.

Fine G.F., Cavanagh L.M., Afonja A., Binions R. 2010. Metal oxide semiconductor gas sensors in environmental monitoring. Sensors. 10:5469-502.

Ghasemi-Varnamkhasti M., Mohtasebi S.S., Rodriguez-Méndez M.L., Lozano J., Razavi S.H., Ahmadi H. 2011. Potential application of electronic nose technology in brewery. Trends. Food. Sci. Tech. 22:165-74.

Giovenzana V., Beghi R., Buratti S., Civelli R., Guidetti R. 2014. Monitoring of fresh-cut valerianella locusta laterr. shelf life by electronic nose and VIS–NIR spectroscopy. Talanta. 120:368-75.

Haugen J.E., Kvaal K. 1998. Electronic nose and artificial neural network. Meat. Sci. 49:273-86.

Karami H., Rasekh M., Mirzaee G.E. 2020. Application of the Enose machine system to detect adulterations in mixed edible oils using chemometrics methods. J. Food. Process. Pres. 44: 14696.

Khodabakhshian R., Emadi B., Khojastehpour M., Golzarian M.R., Sazgarnia A. 2017. Non-destructive evaluation of maturity and quality parameters of pomegranate fruit by visible/near infrared spectroscopy. Int. J. Food. Prop. 20:41-52.

Kodogiannis V.S. 2017 Application of an electronic nose coupled with fuzzy-wavelet network for the detection of meat spoilage. Food. Bioprocess. Tech. 10:730-49.

Mamat M., Samad S.A, Hannan M.A. 2011. An electronic nose for reliable measurement and correct classification of beverages. Sensors. 11:6435-53.

Mavani N.R., Ali J.M., Othman S., Hussain M.A., Hashim H., Rahman N.A. 2021. Application of artificial intelligence in food industry - a guideline. Food. Eng. Rev. 9:1-42.

Pearce T.C. 1997. Computational parallels between the biological olfactory pathway and its analogue the electronic nose': part II. Sensor-based machine olfaction. Biosystems. 41:69-90.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V. 2011. Scikit-learn: machine learning in python. J. Mach. Lear. Res. 12:2825-30.

Poghossian A., Geissler H., Schöning M.J. 2019. Rapid methods and sensors for milk quality monitoring and spoilage detection. Biosens. Bioelectron. 140:111272.

Sanaeifar A., Mohtasebi S.S, Ghasemi-Varnamkhasti M., Ahmadi H., Lozano J. 2014. Development and application of a new low cost electronic nose for the ripeness monitoring of banana using computational techniques (PCA, LDA, SIMCA and SVM). CAAS. Agric. J. 32:538-48.

Schaller E., Bosset J., Escher F. 1998. Electronic noses and their application to food. Food. Sci. Tech-brazil. 31:305-16.

Semwal R., Aier I., Tyagi P., Varadwaj P.K. 2021. DeEPn: a deep neural network based tool for enzyme functional annotation. J. Biomol. Struct. Dyn. 39:2733-43.

Singh B., Chadha K.L., Sahai S. 2010. Performance of litchi cultivar for yield and physico-chemical quality of fruits. Indian. J. Hortic. 67:96-8.

Singh D., Singh B. 2019. Investigating the impact of data normalization on classification performance. Appl. Soft. Comput. 97:105524.

Singh P., Singh I. 1994. Physico-chemical changes during storage of litchi (Litchi chinensis) beverages. Indian. J. Agr. Sci. 64:168-70.

Tan J., Xu J. 2020. Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: a review. Artificial. Intelligence. Agriculture. 4:104-15.

Tang K.T., Chiu S.W., Pan C.H., Hsieh H.Y., Liang Y.S., Liu S.C. 2010. Development of a portable electronic nose system for the detection and classification of fruity odors. Sensors. 10:9179-93.

Voss H.G.J., Stevan J.S.L., Ayub R.A. 2019. Peach growth cycle monitoring using an electronic nose. Comput. Electron. Agr. 163:104858.

Wang L., Brown S.J. 2006. Prediction of DNA-binding residues from sequence features. J. Bioinf. Comput. Biol. 4:1141-58.

Wijaya D.R., Sarno R., Zulaika E., Sabila S.I. 2017. Development of mobile electronic nose for beef quality monitoring. Procedia. Comput. Sci. 124:728-35.

Xu S., Lü E., Lu H., Zhou Z., Wang Y., Yang J., Wang Y. 2016. Quality detection of litchi stored in different environments using an electronic nose. Sensors. 16:852.

Yang X., Chen J., Jia L., Yu W., Wang D., Wei W., Li S., Tian S., Wu D. 2020. Rapid and non-destructive detection of compression damage of yellow peach using an electronic nose and chemometrics. Sensors. 20:1866.

Yang X., Yu Q., He L., Guo T. 2013. The one-against-all partition based binary tree support vector machine algorithms for multi-class classification. Neurocomputing. 113:1-7.

Yoshida K., Ishikawa E., Joshi M., Lechat H., Ayouni F., Bonnefille M. 2012. Quality control and rancidity tendency of nut mix using an electronic nose. In: Indo-Japanese Conference on Perception and Machine Intelligence. Indo-Japanese Conference on Perception and Machine Intelligence. Berlin: Springer, Berlin, Heidelberg. pp. 163-170.

Yu H., Wang J., Yao C., Zhang H., Yu Y. 2008. Quality grade identification of green tea using E-nose by CA and ANN. Food. Sci. Tech-brazil. 41:1268-73.

Zhaoqi Z., Xuequn P., Xuewu D., Zuoliang J. 2002. Change of anthocyanin content and its determination during lychee pericarp browning. J. South. China. Agric. 23:16-9.