Determination of Chopped Fruits Freshness with High Accuracy by Using Electronic Nose

*[1]Bilge Han Tozlu

[1]Department of Electrical Electronics Engineering, Hitit University, Çorum, Türkiye

bilgehantozlu@hitit.edu.tr, ID

Abstract

Especially for the last 20 years many studies have been made in the field of health, chemistry and food with the electronic nose which is a very popular research area thanks to development of sensor technology. In these studies, high achievements have been obtained in many subjects ranging from disease detection to identification of bacterial species and from determination of food quality to aroma separation. In the study, the variation of the freshness of some fruits prepared for eating was examined with an electronic nose according to the days. Fruits as chopped melon, peach, banana and uncut strawberries were put on plates separately and all of them were stored at the room temperature for 5 days. A low cost electronic nose has been made with the MQ branded 11 gas sensors. Fruits were sniffed 15 times for each day at the same time zone. Sensors' data were transferred to the computer via 2 Arduino Uno microcontroller cards and recorded to computer with an interface program made up in the LabVIEW environment. Then, features were extracted from the obtained data taken from the sensors, and the extracted features were classified by using the k-Nearest Neighbors and Neural Network Classification Algorithms. Best classification accuracies were obtained as; 93.28% for melon, 80.80% for peach, 84.80% for banana and 75.78% for strawberry.

Keywords: Electronic nose, fruit freshness, classification

1. INTRODUCTION

Electronic nose (e-nose) is a device that can recognize odors which were introduced previously. An e-nose generally consists of the following parts; (a) an odor box that the odor which is desired to be recognised, will enter, (b) a sensor block which exists in this box, composed of gas sensors that they can detect the amounts of gases in the composition of the incoming odor, (c) an analog-digital converter unit which converts the electrical signals of these electrochemical sensors to digital data, (d) a classification algorithm which extracts features of the collected data and classifies, and (e) a computer where these operations are performed.

Today, many scientists generate e-noses to distinguish and recognize odors in health, food and chemical area for commercial or academic purposes. There have been many successful studies in the field of health. The diseases like lung cancer[1, 2] heart diseases[3], respiratory system diseases[4, 5], Alzheimer's and Parkinson's diseases[6, 7] [5,6] and diabetes[8, 9] were diagnosed with high accuracy from the patients' breath by using an e-nose. Electronic nose studies have been carried out in the cosmetic field in order to differentiate real and fake perfume[10] as well as to determine a special perfume[11].

In the food and beverage area; numerous studies have been done in the evaluation of quality of foods or beverages by using an electronic nose. The quality or aroma of many foods such as tomatoes[12], peaches[13] and beverages such as fruit juice[14], wine[15] and tea[16] have been determined with the e-nose in food industry. Moreover, there are many e-nose studies available on the open literature to determine the shelf life of milk[17] and freshness of food such as fish[18], peach[19], egg[20], etc.

In this study; daily changes of the smell of chopped melon, peach, banana and uncut strawberry have been observed with electronic nose. An electronic nose consisting of 11 gas sensors has been made for this study. The sensor array has been placed in an odor (sample) box. Electrical outputs of the sensors have been converted to digital data via an analog-digital converter and the data have been recorded to the computer. The smells of fruits were recorded for 5 days. Then classification process has been carried out. The output of the study can be useful for commercial kitchens, food factories and electronic kitchen products.

2. MATERIAL AND METHOD

2.1. Customized Electronic Nose Setup

In the study, primarily the sensor block of the electronic nose was generated by using the gas sensors listed in Table1.

Table 1. List of Gas Sensors

Sensor No	Sensor Name	Target Gas
1	MQ-2	Methane, Butane, LPG, Smoke
2	MQ-3	Alcohol, Ethanol, Smoke
3	MQ-4	Methane, CNG Gas
4	MQ-5	Natural Gas, LPG
5	MQ-6	LPG, Butane Gas
6	MQ-7	Carbon Monoxide
7	MQ-8	Hydrogen Gas
8	MQ-9	Carbon Monoxide, Flammable Gasses
9	MQ-131	Ozone
10	MQ-135	Air Quality (CO, Ammonia, Benzene, Alcohol, Smoke)
11	MQ-137	Ammonia

MQ branded gas sensors have been used with their kit in this study. The sensor array has been generated by placing these sensors to 6x10 cm board as shown in Figure 1. In Figure 2, the manufactured electronic nose system is given. The sensor array has been placed inside a storage container and the cables are pulled out of an airtight hole. The supply voltages of the sensor kits have been provided from a power supply and the analog outputs of the sensors have been connected to the analog inputs of 2 Arduino Uno cards. The sensors outputs that are electrical voltages, were converted to digital data by Arduino cards and the data have been transferred to the computer via USB port.

Figure 1. Sensor array
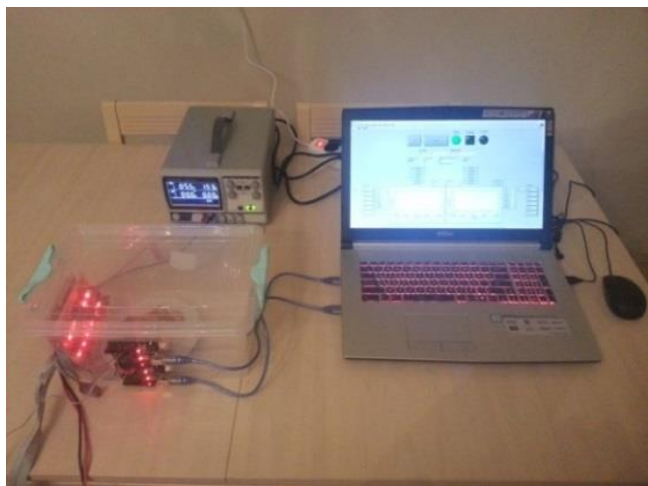
Figure 2. Customized electronic nose system

Sensors' data have been recorded with a data collection software which was prepared in LabVIEW. The presence and quantity of the relevant gases in the smell of the fruits which placed in the sample box, have been sensed by the related sensor and recorded to the computer. And obtained data have been classified after features extracted in MATLAB software.

2.2. Data Collection and Analysis

In this study; strawberries were washed to be eaten 1 day after they were picked and their daily odors were recorded with an e-nose, while other fruits were chopped on the first day they were bought from the market and their odors from chopping to spoiling was studied. The odors of each fruit were taken 15 times in a day along 5 days. All records were taken between 20.00-23.00 hours under conditions of 22-24 °C temperature and 60%-70% humidity. A sniffing cycle consisted of placing the chopped fruit plate into the previously ventilated sample box, closing the lid of the box and starting the odor recording software program. The period of sniffing cycle was determined as 120 seconds and 10 data were taken per second from the sensors. As a result of 1200 data recorded from each of 11 gas sensors, a data matrix which is consisted 11x1200 data were obtained in a sniffing cycle. A 3 dimensional matrix having 15x11x1200 dimensions were obtained daily for each fruit. End of the 5 days, a total of 300 different odors were recorded, 75 for each fruit.

For the classification process, firstly, Kurtosis, Skewness, Sum, Average, Hilbert Transformation and Variance of the Derivative (VD) methods were applied to 11 sensors' data of each fruit's 75 odor records and the features of these data were extracted. How these features are calculated[21] is given in Table 2.

In the formulas given in Table 2, x_i is the ith sample of a trial x, $\overline{x}$ is the mean value of x, x' is the derivative of x, x'_i is the ith sample of x', $\overline{x'}$ is the mean value of x', L is the length of a trial, a_{mean} and a_{std} are the mean and standard deviation of real part of $\hat{s}(t)$ respectively.

The kNN classification method is an extremely useful method that is frequently used in classification problems. In kNN classification, unknown specimen is classified according to nearest k neighbours. In this method; the metric distances of the training trials to the test trial are calculated, after that nearest neighbors in k quantity are taken into account and the test trial is labeled to be the class of the largest number of these training data. Ideal k number was determined with random sub-sampling method[21].

The neural network classification is also a method that is frequently used in classification studies. In this method; the data in the input layer is processed with the activation functions in the neurons in the hidden layer or layers, and the results of neurons in hidden layer transferred to the output layer. The transitions of the data between the layers are made according to certain weights. Finally, the data coming to the

output layer with the determined weights are assigned to a class as a result of the evaluation made in the output layer. In this study, a single layer is used in the hidden layer.

Table 2. Methods of Extracting Features

Kurtosis	$\dfrac{\dfrac{1}{L}\sum_{i=1}^{L}(x_i-\overline{x})}{\left(\dfrac{1}{L}\sum_{i=1}^{L}(x_i-\overline{x})^2\right)^2}$
Skewness	$\dfrac{\dfrac{1}{L}\sum_{i=1}^{L}(x_i-\overline{x})^3}{\left(\sqrt{\dfrac{1}{L}\sum_{i=1}^{L}(x_i-\overline{x})^2}\right)^3}$
Sum	$\sum_{i=1}^{L}(x_i)$
Mean	$\dfrac{1}{L}\sum_{i=1}^{L}(x_i)$
VD	$\dfrac{1}{L}\sum_{i=1}^{L}(x'_i-\overline{x'})^2$
Hilbert Transform	$H\{s(t)\}=\hat{s}(t)=s(t)*\dfrac{1}{\pi t}$ $=\dfrac{1}{\pi}\int_{-\infty}^{+\infty}\dfrac{s(\tau)}{t-\tau}d\tau=a+jb$ $a_{mean}=\sum_{a=1}^{L}\dfrac{a_1+a_2+...+a_L}{L}$ $a_{std}=\sqrt{\dfrac{(a_1-a_{mean})^2+...+(a_L-a_{mean})^2}{L-1}}$

It was used classification accuracy (CA), sensitivity (SE) and specificity (SF) metrics for evaluating performance of the classifiers. CA is the percentage expression of the ratio of the number of correctly classified trials to the total number of trials (Eq. 1). SE and SF were also calculated as follows (Eq. 2 anad Eq. 3):

$$CA=\frac{CCT}{TT}\,x100 \tag{1}$$

$$SE=\frac{TP}{TP+FN}\,x100 \tag{2}$$

$$SF=\frac{TN}{TN+FP}\,x100 \tag{3}$$

CCT, TT, TP, TN, FP, FN are abbreviations of; the number of correctly classified trials, the total number of considered trials, the number of positive samples correctly classified, the number of negative samples correctly classified, the number of positive samples incorrectly classified, the number of negative samples incorrectly classified, respectively.

Features are extracted from the training data set. By using all features, the feature that gives the highest cross validation accuracy (CVA) is determined. Then, the remaining features are added next to this feature by trying all combinations, and the features used for the highest CVA are recorded. This process is done as much as the determined number (100 times for this study) of cross validations with randomly selected sub-training and validation sets. In the cross validation process given in Figure 3, the most used feature and the features which were used more than half of the highest usage count, are selected as the effective features of the classification.

The classification process was carried out in 5 classes for 5 different days. Randomly selected 10 of the 15 data which was recorded in each day, used as training data and the remaining 5 were used as test data. 5 of the 10 training data were used as sub-learning set by selecting randomly, and the other 5 were used as validation data. Feature selection was done as described above. Then, using the determined features, 5 data allocated for the test were classified by kNN and NN classification algorithms.

Thus, one classification process was completed and classification accuracy, sensitivity and specificity were recorded. In order to increase the reliability of the classification accuracy, the classification process mentioned above was performed 100 times for different training-test sets. The arithmetic mean of the results of 100 classifications was accepted as the classification result.

3. RESULTS

In this study, four fruits' five days odors have been classified by using an electronic nose. Totally 300 odors 75 for each fruit have been sniffed, and gas sensors' data has been recorded to computer. All sniffing cycles composed of 2 minutes records. 1200 values for each of 11 gas sensors have been got in a sniffing cycle. At the classification stage; 50 of 75 data were used as training data (10 of the 15 per day) and the other 25 of them were classified to 5 classes with kNN and NN classification algorithms. Classification flow diagram is given in Figure 3.

The classification results for each fruit are as follows:

For melon;

An average of 93.28% classification accuracy (CA), 85.40% sensitivity (SE) and 97.00% specificity (SF) were achieved at the end of 100 classification cycles with kNN algorithm. By using NN algorithm, 84.53% CA, 66.60% SE and 97.45% SF were achieved at the end of 100 classification cycles. Most effective feature is x_{mean} of Hilbert Transform of MQ-4 gas sensor. Error matrix of a randomly selected

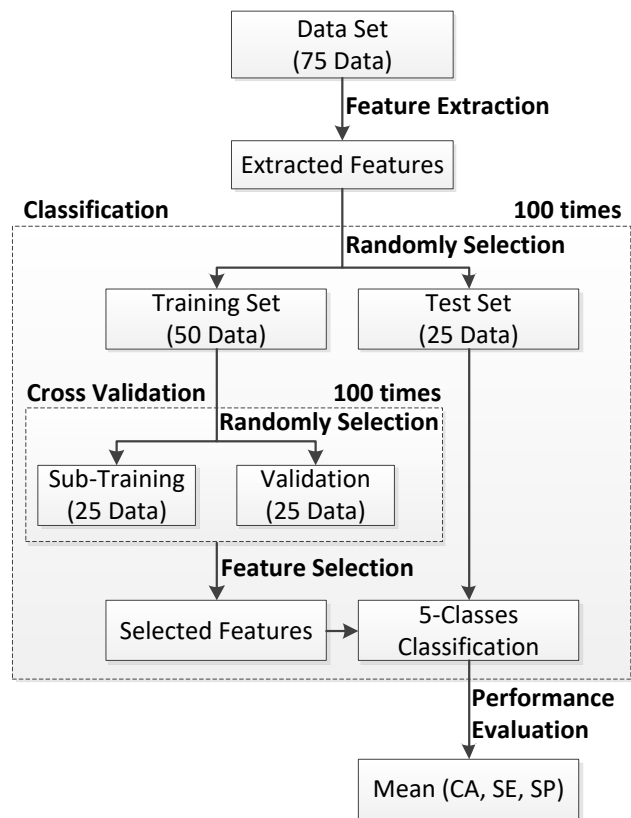classification among 100 different classifications is given in Table 3.

Figure 3. Classification Flow Diagram

Table 3. Error Matrix

		Predicted				
		1.Day	2. Day	3. Day	4. Day	5. Day
Actual	1. Day	4	0	0	0	0
	2. Day	0	5	0	0	0
	3. Day	1	0	5	0	0
	4. Day	0	0	0	4	0
	5. Day	0	0	0	1	5

In Figure 4, extracted features graph is given for random one cycle of 100 cycles. Symbols in given Figure 4-7 represent the days as follow:

+ : 1. Day   □ : 3. Day   □ : 5. Day
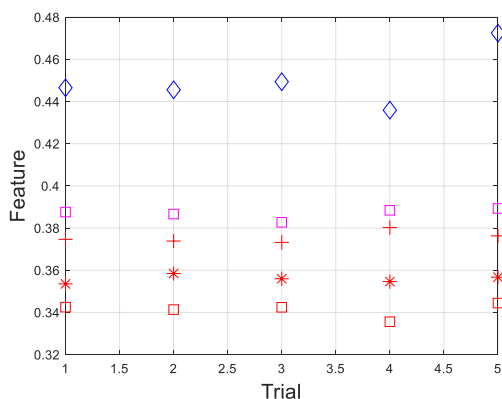◇ : 2. Day   * : 4. Day

Figure 4. The graph of extracted feature of the chopped melon

For peach;

An average of 80.80% CA, 95.00% SE and 93.75% SF were achieved with kNN algorithm and 76.52% CA, 89.00% SE and 91.05% SF were achieved by using NN algorithm at the end of 100 classification cycles. Most effective feature is x_{mean} of Hilbert Transform of MQ-3 gas sensor's data. Error matrix of a randomly selected classification among 100 different classifications is given in Table 4. In Figure 5, extracted feature graph is given for one classification.

Table 4. Error Matrix

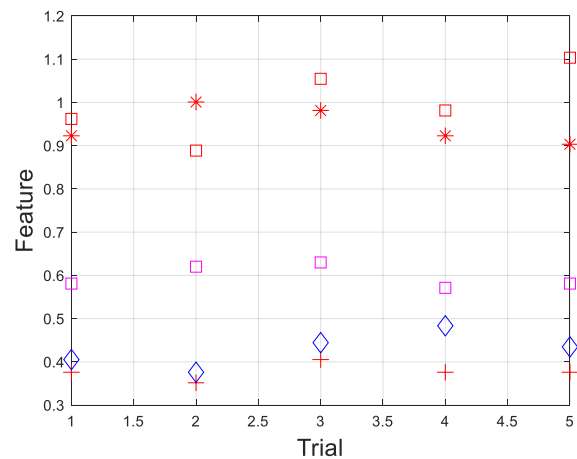|		Predicted				
		1.Day	2. Day	3. Day	4. Day	5. Day
Actual	1. Day	5	1	0	0	0
	2. Day	0	4	0	0	0
	3. Day	1	0	5	0	0
	4. Day	0	0	0	2	1
	5. Day	0	0	0	3	4

Figure 5. The graph of extracted feature of the chopped peach

For banana;

An average of 84.80% CA, 83.40% SE, 96.80% SF were achieved with kNN algorithm, and 75.79% CA, 83.20% SE, 94.95% SF were achieved with NN algorithm at the end of 100 classification cycles. Most effective features are x_{mean} of Hilbert Transform of MQ-4 gas sensor's data and x_{std} of Hilbert Transform of MQ-131 gas sensor's data. Error matrix of one among 100 classifications is given in Table 5. Extracted feature graph for one of the classifications is given in Figure 6.

Table 5. Error Matrix

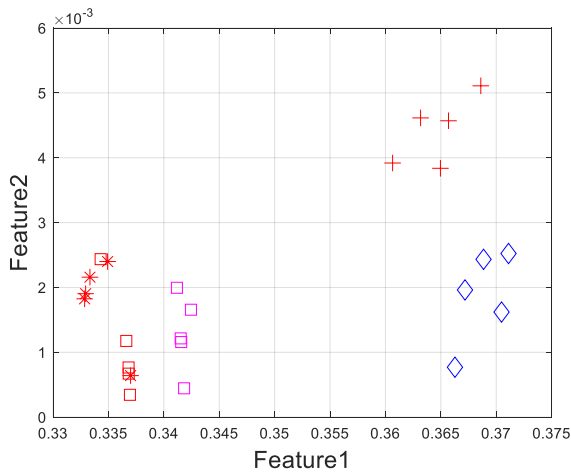|		Predicted				
		1.Day	2. Day	3. Day	4. Day	5. Day
Actual	1. Day	3	0	0	0	0
	2. Day	2	5	0	0	0
	3. Day	1	0	5	0	0
	4. Day	0	0	0	3	0
	5. Day	0	0	0	2	5

Figure 6. The graph of extracted features of the chopped banana

For strawberry;

An average of 75.78% CA, 92.80% SE, 96.00% SF were achieved with NN algorithm, and 65.72% CA, 93.00% SE, 97.90% SF were achieved with kNN algorithm at the end of 100 classification cycles. Most effective features are x_{mean} and x_{std} of Hilbert Transform of MQ-2, MQ-3 and MQ-4 gas sensors' data. Error matrix of one among 100 classifications is given in Table 6. Extracted feature graph for one of the classifications is given in Figure 7.

Table 6. Error Matrix

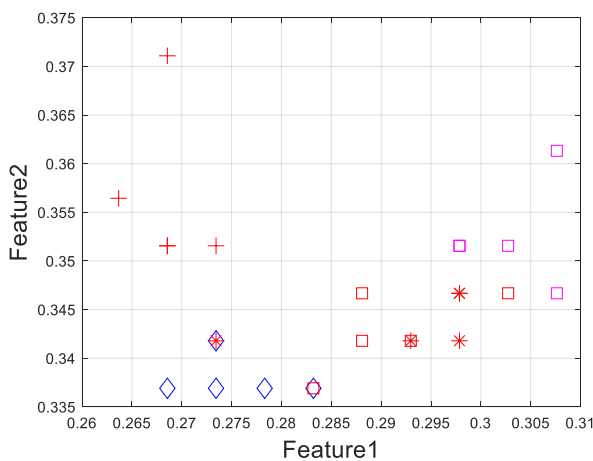|		Predicted				
		1.Day	2. Day	3. Day	4. Day	5. Day
Actual	1. Day	5	0	0	0	0
	2. Day	0	4	0	0	0
	3. Day	0	0	2	0	0
	4. Day	0	1	3	3	2
	5. Day	0	0	0	2	3

Figure 7. The graph of extracted features of the strawberry

4. DISCUSSION AND CONCLUSION

In this study, the freshness of the fruits was determined with high accuracy with an electronic nose. To start with, a low cost electronic nose was made up. Then, melon, banana and peach were chopped to plates separately and unchopped

strawberries placed to another plate, and these fruits were kept in room temperature. These fruits' odors were sniffed with the e-nose in the same time zone (between 20.00-23.00 hours) for 5 days. A total of 75 odor data for each fruit, 15 per day, were recorded into the computer. Finally, 25 of data were classified with two classifier algorithms by introducing randomly selected 50 of data as training data, after extracting features. The success rates and related classifier types of classifications, which are more successful are given in Table 7.

Table 7. Classifications Summary

Fruit	CA (%)	SE (%)	SF (%)	Classification Algorithm
Melon	93.28	85.40	97.00	kNN
Peach	80.80	95.00	93.75	kNN
Banana	84.80	83.40	96.80	kNN
Strawberry	75.78	92.80	96.00	NN

While the kNN classification algorithm can detect freshness and spoilage in melon, peach and banana fruits with high accuracy, it was not very successful in strawberry. Strawberry classification could be done with NN with less success than other fruits. The reason of the low classification performance in strawberry is considered to be that it is not chopped. As a matter of fact, it has been scientifically proven that chopped fruits will lose their freshness faster and spoil earlier[22]. It is predicted that strawberry freshness determination will be made with higher accuracy with e-nose by using chopped strawberries in the new studies in this field.

The information of how many days ago the fruit was chopped can be useful for commercial kitchens, food factories and perhaps smart refrigerators to be developed. So, the study has to be developed by increasing the variety of fruits and vegetables. In addition, in scientific studies, it has been suggested to use electronic gas sensors instead of commercial mass spectrometers or gas chromatography in detecting the freshness of fruits [22]. Undoubtedly, the odors emitted by the fruits change daily and these can be detected by the gas chromatography method. However, the gas chromatography method is quite costly, laborious and long-term work compared to the electronic nose method. With the electronic nose, a situation can be detected so quickly from an odor, and this method can be integrated into industry very easily.

This study has some limitations: The low number of data is one of the negative aspects of the study. Although high accuracy classification results show that the spoilage of fruits can be monitored daily by their fragrances, working with larger data sets in future studies will further reinforce the reliability of the method. The other limitation is lack of gas chromatography of fruit spoiling before this study for choosing related gas sensors.

Presentation at a meeting: A part of this study has been shared as an oral presentation at 3rd International Conference on Advanced Engineering Technologies (ICADET) in Bayburt, TURKEY on 19-21 September 2019

REFERENCES

[1] A. D'Amico vd., "An investigation on electronic nose diagnosis of lung cancer", Lung Cancer, c. 68, sy 2, ss. 170-176, May. 2010, doi: 10.1016/j.lungcan.2009.11.003.

[2] R. F. Machado vd., "Detection of Lung Cancer by Sensor Array Analyses of Exhaled Breath", Am J Respir Crit Care Med, c. 171, sy 11, ss. 1286-1291, Haz. 2005, doi: 10.1164/rccm.200409-1184OC.

[3] B. H. Tozlu, C. Şimşek, O. Aydemir, ve Y. Karavelioglu, "A High performance electronic nose system for the recognition of myocardial infarction and coronary artery diseases", Biomedical Signal Processing and Control, c. 64, s. 102247, Şub. 2021, doi: 10.1016/j.bspc.2020.102247.

[4] S. Scarlata, G. Pennazza, M. Santonico, C. Pedone, ve R. A. Incalzi, "Exhaled breath analysis by electronic nose in respiratory diseases", Expert Review of Molecular Diagnostics, c. 15, sy 7, ss. 933-956, Tem. 2015, doi: 10.1586/14737159.2015.1043895.

[5] N. Fens vd., "Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma", Am. J. Respir. Crit. Care Med., c. 180, sy 11, ss. 1076-1082, Ara. 2009, doi: 10.1164/rccm.200906-0939OC.

[6] J.-P. Bach vd., "Measuring Compounds in Exhaled Air to Detect Alzheimer's Disease and Parkinson's Disease", PLOS ONE, c. 10, sy 7, s. e0132227, Tem. 2015, doi: 10.1371/journal.pone.0132227.

[7] U. Tisch vd., "Detection of Alzheimer's and Parkinson's disease from exhaled breath using nanomaterial-based sensors", Nanomedicine, c. 8, sy 1, ss. 43-56, Eki. 2012, doi: 10.2217/nnm.12.105.

[8] S. Esfahani, A. Wicaksono, E. Mozdiak, R. P. Arasaradnam, ve J. A. Covington, "Non-Invasive Diagnosis of Diabetes by Volatile Organic Compounds in Urine Using FAIMS and Fox4000 Electronic Nose", Biosensors (Basel), c. 8, sy 4, Ara. 2018, doi: 10.3390/bios8040121.

[9] A. Bermak ve M. Hassan, "Noninvasive Diabetes Monitoring with Electronic Nose", Mar. 2016, c. 2016, s. HBPP2776. doi: 10.5339/qfarc.2016.HBPP2776.

[10] J. Gebicki, B. Szulczynski, ve M. Kaminski, "Determination of authenticity of brand perfume using electronic nose prototypes", Meas. Sci. Technol., c. 26, sy 12, s. 125103, Eki. 2015, doi: 10.1088/0957-0233/26/12/125103.

[11] A. Carrasco, C. Saby, ve P. Bernadet, "Discrimination of Yves Saint Laurent perfumes by an electronic nose", Flavour and Fragrance Journal, c. 13, sy 5, ss. 335-348, Eyl. 1998.

[12] X. Huang, S. Pan, Z. Sun, Y. Wei-tao, ve J. H. Aheto, "Evaluating quality of tomato during storage using fusion information of computer vision and electronic nose", Ağu. 2018, [Çevrimiçi]. Erişim adresi: https://doi.org/10.1111/jfpe.12832

[13] "Evaluation of peach quality indices using an electronic nose by MLR, QPST and BP network", Sensors and Actuators B: Chemical, c. 134, sy 1, ss. 332-338, Ağu. 2008, doi: 10.1016/j.snb.2008.05.008.

[14] "Qualification and quantisation of processed strawberry juice based on electronic nose and tongue", LWT - Food Science and Technology, c. 60, sy 1, ss. 115-123, Oca. 2015, doi: 10.1016/j.lwt.2014.08.041.

[15] M. Aleixandre, J. M. Cabellos, T. Arroyo, ve M. C. Horrillo, "Quantification of Wine Mixtures with an Electronic Nose and a Human Panel", Front. Bioeng. Biotechnol., c. 6, 2018, doi: 10.3389/fbioe.2018.00014.

[16] B. Tozlu, H. I. Okumus, ve C. Simsek, "Onlıne Qualıty Classıfyıng Wıth Electronıc Nose For Black Tea Productıon.", International Journal of Academic Research, c. 6, sy 4, 2014.

[17] S. Labreche, S. Bazzo, S. Cade, ve E. Chanie, "Shelf life determination by electronic nose: application to milk", Sensors and Actuators B: Chemical, c. 106, sy 1, ss. 199-206, Nis. 2005, doi: 10.1016/j.snb.2004.06.027.

[18] S. Güney ve A. Atasoy, "Freshness Classification of Horse Mackerels with E-Nose System Using Hybrid Binary Decision Tree Structure", Int. J. Patt. Recogn. Artif. Intell., c. 34, sy 03, s. 2050003, May. 2019, doi: 10.1142/S0218001420500032.

[19] "Study of peach freshness predictive method based on electronic nose", Food Control, c. 28, sy 1, ss. 25-32, Kas. 2012, doi: 10.1016/j.foodcont.2012.04.025.

[20] R. Dutta, E. L. Hines, J. W. Gardner, D. D. Udrea, ve P. Boilot, "Non-destructive egg freshness determination: an electronic nose based approach", Meas. Sci. Technol., c. 14, sy 2, ss. 190-198, Oca. 2003, doi: 10.1088/0957-0233/14/2/306.

[21] E. Ergün ve Ö. Aydemir, "Decoding of Binary Mental Arithmetic Based Near-Infrared Spectroscopy Signals", içinde 2018 3rd International Conference on Computer Science and Engineering (UBMK), Eyl. 2018, ss. 201-204. doi: 10.1109/UBMK.2018.8566462.

[22] A. Ceccarelli vd., "Nectarine volatilome response to fresh-cutting and storage", Postharvest Biology and Technology, c. 159, s. 111020, Oca. 2020, doi: 10.1016/j.postharvbio.2019.111020.