Monitoring and Detection of Fruits and Vegetables Spoilage in the Refrigerator using Electronic Nose Based on Principal Component Analysis

Meo Vincent C. Caya[1,2]; Febus Reidj G. Cruz[1,2]; Carolle Marian N. Fernando[1]; Roselle Marian M. Lafuente[1]; Mark B. Malonzo[1]; Wen-Yaw Chung[2,*]

[1]*School of Electrical, Electronics and Computer Engineering*
Mapua University
Manila, Philippines

[2]*Department of Electronic Engineering*
Chung-Yuan Christian University
Chungli, Taiwan

[*] eldanny@cycu.edu.tw

***Abstract*—The study focuses on monitoring freshness and determining food spoilage inside the refrigerator. The objective is to design an electronic nose system that will be sensitive to the gases emitted by spoiled food samples namely banana, pechay, carrots and grapes operating in low level temperature particularly the refrigerator and then determine food spoilage using Principal Component Analysis – K Nearest Neighbors, however, it will not take any corrective actions. The system will gather readings from MQ gas sensors and will be subjected to PCA and KNN. PCA is implemented to minimize data and for feature projection represented in form of graphs. Whereas, KNN is applied for clusters formed by the PCA transformation to classify the grouping of the food. The results from the combined approach produced an overall accuracy rate of 92%, thus, the electronic nose system is capable of sensing food gases and accurately determine spoilage inside the refrigerator.**

Index Terms*—*Electronic Nose, Principal Component Analysis, K-nearest Neighbor, Classification, MQ Sensors

I. INTRODUCTION

Electronic nose is a system that functions like a human olfactory system. It is used to analyze a set of chemicals that a certain object emits. The main components of an e-nose are: gas sensors array and an algorithm to identify the pattern of the results. First, the array of sensors will detect the odor molecules, and then the signals produced are extracted to feed the algorithm for classification [1]. The general architecture of E-nose comprises of sampling chamber for gas that interacts to sensors chamber and uses a software for data acquisition and analyzer to produce an output as shown in Fig. 1 [2].

There are several e-nose devices applications such as food processing, medical and gas concentration analysis. E-nose can be integrated in consumer electronics such as chiller and refrigerator for development of smart packaging system [3]. Implementing e-nose on a smart home has been used for identifying hazards in ambient-assisted living environments, particularly the spoilage of food inside the home [4].

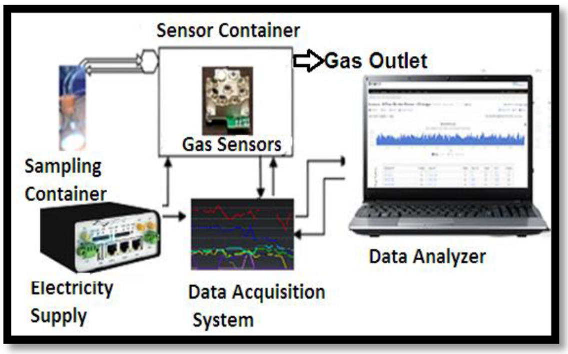

Figure 1 E-nose System Architecture [2]

E-nose is commonly used in food industry, for example discriminating chicken freshness during storage based on PCA method [5], monitoring of anchovy freshness using sensor array and based on pattern recognition techniques: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) [6], Classification of Beef and Pork Aroma Based using PCA [1] [7], meat authentication [8], meat and fish freshness [9], food freshness [2], fruit identification, ripeness assessment and assuring food quality [10], fruit ripeness determination [11] [12] e.g. cocoa ripeness classification [13], banana shelf life [14] and mango grading [15] and sorting many more. Thus, freshness of food is really important in order to avoid health risks especially in schools, hospitals and markets where food safety incidents often occur. The determination of freshness and quality of food is an essential way to sustain quality of nutrition, to prevent food poisoning and wastage and the like. Foods can be classified as good or bad based on the gases produced from microorganisms' activity. The higher the microorganism content, the higher chance of food spoilage. There are many methods in determining the quality of the food such as the visual appearance, taste, and smell. However, this method is subjective as human sense may vary from each person. Thus, e-nose is considered as an effective way to simulate human olfactory system [2].

978-1-7281-3044-6/19/$31.00 ©2019 IEEE

The common classification technique used for electronic nose applications is Principal Component Analysis (PCA). PCA is a mathematical tool for analyzing, classifying and minimizing datasets in a multivariate problem while retaining most information. The coordinates of the data (scores plot) is used for classification of data cluster [6].

In this study, the researchers aim to develop an electronic nose system that can monitor freshness, detect food spoilage and can operate on a low-level temperature via Principal Component Analysis (PCA). Specifically, this work aimed to: (1) design an electronic nose system that will be sensitive to the organic volatile gases from spoilage of food in low level temperatures; (2) to examine and detect the quality and production of gases of fruits and vegetables using the electronic nose system.

II. METHODOLOGY

A. Block Diagram

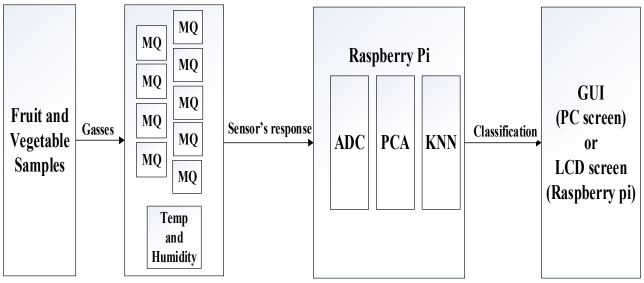

Figure 2 E-nose System for Spoilage Detection Block Diagram

Figure 2 represents the block diagram of the e-nose system to be used to detect the volatile gases that spoiling fruits or vegetables emit inside the refrigerator. The e-nose system comprises of (1) the sensing mechanism that will generate voltages depending on the gases it will detect and (2) the Raspberry Pi will transform the signal generated from the sensors to binary values through ADC, process the data using PCA and classify these data using kNN algorithm. The output classification will then be shown on the connected display screen on the Raspberry Pi. The input comes from the Fruit and Vegetable samples. The gases emitted from the samples is sensed by the array of sensors made up of MQ sensors; (Sulfur Dioxide; Methane, Alcohol, Smoke; Ammonia; Alcohol; Combustible Gases, LPG, Propane, Hydrogen; Ammonia, Sulfide, Benzene; Ozone; Combustible Gases, Natural Gas; and Carbon Monoxide, Methane, Propane) a series of gas-sensitive sensors that detects volatile compounds concentrations. The concentration of the gases present inside the refrigerator generate a collective response that is sent to the Raspberry Pi analog input ports. The collected data from the sensors array will be used to feed the Principal Component Analysis Algorithm. PCA eliminates unwanted data or noise. It is a type of Data Reduction technique used for analyzing a collective response such as gas concentrations. Afterwards, the result of the PCA is then classified using kNN algorithm. The e-nose system basically monitors the gases emitted by fruits and vegetables present inside the refrigerator then once food spoilage is detected a warning will be displayed on the LCD screen.

B. Software Development

Principal Component Analysis algorithm is a technique commonly used for large scale data set and it is often used for applications such as data compression and blind source separation. Data compression is one of the applications of PCA since it is capable of summarizing large data sets into common points where some variables can be associated to another which is also referred to as principal components. The PCA is also represented as plots of data when translated into graphical results [16]. PCA is implemented on normalized data using sklearn library from pycharm software to minimize, scale and center the gathered data.

The classifier used to categorize the input fruit and vegetable samples is KNN algorithm. This method will get the K most similar neighbors from PCA. The values gathered from fresh and spoiled database will undergo elective process on which group has the most elements that are nearest to the new input's location in the plot. The system makes use of two classes: class 1 (Spoiled) and class 0 (Not Spoiled). Fig. 3 shows plotted PCA points for spoiled samples. The samples for each class are determined by the middle values of each possible humidity gathered from training. Before plotting the data points, the array of values is pre-processed by normalizing rows and columns into scaled values. This process transforms and fits the data points to the center of the plot and contains them in a certain range of values along the x and y axes. This is necessary due to differing scales in readings between sensors. The value of k for the k-NN algorithm indicates the number of neighbors that are needed to classify a certain point. The test data set will be identified as either: Spoiled and Not Spoiled. The k-values used is odd since our class category is even and to avoid having a 50-50 result. The addition of the k-NN algorithm is then applied to prepare for testing. This is done by using a query point or a point of reference where the k-nearest neighbors from the data points will be determined.

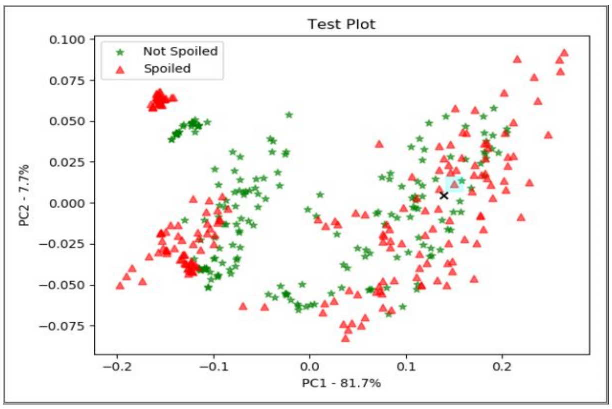

Figure 3 Spoiled sample query point classified as spoiled

C. Hardware Development

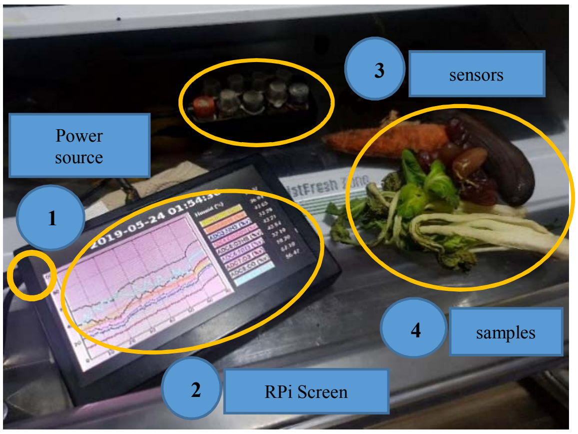

Figure 4 E-nose prototype Actual Setup Inside Refrigerator

The Figure 4 shown above is the actual setup of electronic nose placed inside the refrigerator. To develop the electronic nose system for spoilage detection, the main hardware components used are: (1) Raspberry Pi 3 and (2) Array of Gas Sensors.

The Raspberry Pi (RPi) will be the system's computing device that collects the signals from the gas sensors and sends results to LCD as an output display. The sensors and LCD are connected to the raspberry pi through the USB ports and GPIO pins. The Raspberry pi power source is 5V via micro USB or the GPIO header [17]. To setup the RPi, we will connect it to a PC or laptop for programming purposes. The software used is PyCharm for python programming. During monitoring, LCD screen will display real-time graph of sensors readings, sensors target gasses, sensors values, humidity and the current date and time. When spoilage is detected, a warning message is displayed. The experimental setup focused only on the gases emitted by banana, grapes, carrot, and pechay and it is placed inside the refrigerator crisper drawer because this would serve as a chamber for our samples. The air conditioning of the refrigerator would keep the air regulated or it is for better gas flow of the emitted gas samples. We have observed that it's much better to leave the refrigerator door closed at all times as this provides better reading results.

The sensing element of the electronic nose system is made up of MQ series sensors. The array of nine different gas sensors is implemented because it is more efficient as it detects wider range of gas concentrations. Although, the response of each sensor is unique, but it can be learned through system training to recognize the different patterns [18]. Since the sensor array will be affected by environmental changes inside the refrigerator, an Si7006 sensor will be included to monitor temperature and humidity. The gas sensors array detects various gases which are emitted from the fruits and vegetables samples but the typical gases are alcohol, methane (CH_4), carbon monoxide (CO), ammonia (NH_3) and carbon dioxide (CO_2) [19]. The system used the following sensors to detect the presence of particular gases: MQ-136 to detect Sulfur Dioxide (SO_2), MQ-4 to detect Methane (CH_4), MQ-137 to detect Ammonia (NH_3), MQ-3 for Alcohol (C_2H_5OH), MQ-2 for Propane (C_3H_8), MQ-135 for Ammonia (NH_3), MQ-131 for Ozone (O_3). MQ8- for Combustible Gases (CO), and MQ-9 for Carbon Monoxide (CO). This gas sensor array will be set inside a gas chamber or in this case, the refrigerator which is similar to an airtight food grade box to collect all gases emitted from the samples.

III. Results and Discussion

The researchers gathered total of four samples: banana, pechay, carrot and grapes for training and testing the system. Data acquired from the samples during system training will be used as a reference for testing the system to classify the sample whether spoiled or not spoiled. The wide range training data we have covered all controlled and uncontrolled readings but taking into consideration readings with relative humidity where value ranges from 2-22 since it is the measured humidity range inside the refrigerator.

TABLE I
TESTING DATA RESULTS FOR SPOILED PECHAY

Trial	Percentage Classification (%) at $k = 1$	Human Classification	E-nose Classification
1	100.00	Spoiled	Spoiled
2	100.00	Spoiled	Spoiled
3	100.00	Spoiled	Spoiled
4	100.00	Spoiled	Spoiled
5	100.00	Spoiled	Spoiled
6	100.00	Spoiled	Spoiled
7	100.00	Spoiled	Spoiled
8	100.00	Spoiled	Spoiled
9	91.67	Spoiled	Spoiled
10	91.67	Spoiled	Spoiled
11	41.67	Spoiled	Not Spoiled
12	91.67	Spoiled	Spoiled
13	25.00	Spoiled	Not Spoiled
14	91.67	Spoiled	Spoiled
15	41.67	Spoiled	Not Spoiled

Table I shows the prediction accuracy at k=1, the comparison of human and e-nose classification. The prediction accuracy means the rate of how correctly the sample was predicted. For this test, the only spoiled sample tested is pechay. Out of 15 performed tests, 12 were correctly identified.

TABLE II
TESTING DATA RESULTS FOR SPOILED PECHAY AND CARROT

Trial	Percentage Classification (%) at $k = 1$	Human Classification	E-nose Classification
1	33.33	Spoiled	Not Spoiled
2	100.00	Spoiled	Spoiled
3	100.00	Spoiled	Spoiled
4	83.33	Spoiled	Spoiled
5	91.67	Spoiled	Spoiled
6	58.33	Spoiled	Spoiled
7	75.00	Spoiled	Spoiled
8	100.00	Spoiled	Spoiled
9	100.00	Spoiled	Spoiled
10	100.00	Spoiled	Spoiled
11	100.00	Spoiled	Spoiled
12	100.00	Spoiled	Spoiled
13	100.00	Spoiled	Spoiled
14	100.00	Spoiled	Spoiled
15	100.00	Spoiled	Spoiled

Table II shows the testing results for spoiled pechay and carrot. 14 out of 15 spoiled pechay and carrot were correctly identified.

TABLE III
TESTING DATA RESULTS FOR SPOILED PECHAY, CARROT AND GRAPES

Trial	Percentage Classification (%) at $k = 1$	Human Classification	E-nose Classification
1	58.33	Spoiled	Spoiled
2	83.33	Spoiled	Spoiled
3	58.33	Spoiled	Spoiled
4	75.00	Spoiled	Spoiled
5	100.00	Spoiled	Spoiled
6	75.00	Spoiled	Spoiled
7	41.67	Spoiled	Not Spoiled
8	91.67	Spoiled	Spoiled
9	100.00	Spoiled	Spoiled
10	83.33	Spoiled	Spoiled
11	100.00	Spoiled	Spoiled
12	100.00	Spoiled	Spoiled
13	100.00	Spoiled	Spoiled
14	100.00	Spoiled	Spoiled
15	100.00	Spoiled	Spoiled

Table III shows that the system correctly predicted 14 tests for spoiled pechay, carrot and grapes out of 15 tests performed. Table IV and V show the test results for “all spoiled” samples and all “not spoiled” samples.

TABLE IV
TESTING DATA RESULTS FOR ALL SPOILED SAMPLES

Trial	Percentage Classification (%) at $k = 1$	Human Classification	E-nose Classification
1	91.67	Spoiled	Spoiled
2	100.00	Spoiled	Spoiled
3	100.00	Spoiled	Spoiled
4	100.00	Spoiled	Spoiled
5	100.00	Spoiled	Spoiled
6	100.00	Spoiled	Spoiled
7	100.00	Spoiled	Spoiled
8	75.00	Spoiled	Spoiled
9	100.00	Spoiled	Spoiled
10	91.67	Spoiled	Spoiled
11	100.00	Spoiled	Spoiled
12	100.00	Spoiled	Spoiled
13	100.00	Spoiled	Spoiled
14	100.00	Spoiled	Spoiled
15	100.00	Spoiled	Spoiled

Table IV shows that all tests for All Spoiled samples were correctly identified by the system.

TABLE V
TESTING DATA RESULTS FOR NOT SPOILED SAMPLES

Trial	Percentage Classification (%) at $k = 1$	Human Classification	E-nose Classification
1	100.00	Not Spoiled	Not Spoiled
2	100.00	Not Spoiled	Not Spoiled
3	100.00	Not Spoiled	Not Spoiled
4	100.00	Not Spoiled	Not Spoiled
5	100.00	Not Spoiled	Not Spoiled
6	100.00	Not Spoiled	Not Spoiled
7	100.00	Not Spoiled	Not Spoiled
8	100.00	Not Spoiled	Not Spoiled
9	100.00	Not Spoiled	Not Spoiled
10	100.00	Not Spoiled	Not Spoiled
11	83.33	Not Spoiled	Not Spoiled
12	91.67	Not Spoiled	Not Spoiled
13	33.33	Not Spoiled	Spoiled
14	66.67	Not Spoiled	Not Spoiled
15	75.00	Not Spoiled	Not Spoiled

Table V shows the testing results for all not spoiled samples. Based on the table, 14 out of 15 were correctly predicted.

The results of the k-nn can be analyzed using the confusion matrix as shown on Table VI. The predicted classification is displayed by columns and actual classification is displayed by rows. The diagonal pattern implies true positives or the correct prediction during system testing [20]. The five classes for predicted and actual are the following: spoiled pechay, spoiled pechay and carrot, spoiled pechay, carrot and grapes, all spoiled, and not spoiled and the total predictions made for the classifier is 75 since we consider 15 tests for each of the class.

TABLE VI
CONFUSION MATRIX BASED ON TEST DATA RESULTS

		Predicted				
		Spoiled Pechay	**Spoiled Pechay Carrot**	**Spoiled Pechay Carrot Grapes**	**All Spoiled**	**Not spoiled**
	Spoiled Pechay	12	0	0	0	0
	Spoiled Pechay Carrot	0	14	0	0	0
	Spoiled Pechay Carrot Grapes	0	0	14	0	0
Actual	**All Spoiled**	0	0	0	15	1
	Not Spoiled	3	1	1	0	14

Table VI shows the confusion matrix of the spoilage detection system using electronic nose based on the test data results. For “all spoiled” samples set, 15 out of 15 were correctly identified while for “not spoiled” samples set 14 out of 15 were correctly identified. It is observed that the number of spoiled samples is a factor for the prediction of the system. The greater the number of spoiled samples, the better the prediction accuracy.

The overall accuracy of the system is computed using equation 1. True positive means the summation of the correct prediction during system testing for the following classifications: spoiled pechay, spoiled pechay and carrot, spoiled pechay, carrot, and grapes, all spoiled and not spoiled. Hence, to obtain the overall system accuracy, the calculated true positive of the system is then divided to the total number of tests and then multiplied by 100.

$$Accuracy = \frac{true\ \ positive}{total\ number\ of\ test} \times 100\ (Eq. 1)$$

$$Accuracy = \frac{12+14+14+15+14}{15+15+15+15+15} \times 100 = 92\%$$

Based on confusion matrix results, the calculated accuracy rates per classification are the following: 80% for spoiled pechay, 93.33% for spoiled pechay and carrots, 93.33% for spoiled pechay, carrots and grapes, 100% for all spoiled samples and 93.33% for not spoiled samples. The overall accuracy of the system is 92% which could prove that it is a reliable system for detecting spoilage of food inside the refrigerator. Based on test results, better accuracy is observed when the classification is done per hour and when testing is done in controlled environment although in this paper we have covered both controlled and uncontrolled environments for better accuracy.

IV. CONCLUSION

In this study, the researchers were able to develop an electronic nose system that can determine if the fruits and vegetables inside the refrigerator is spoiled or not spoiled. The system can classify spoilage by using PCA Algorithm and KNN classifier.

The electronic nose system has achieved an overall accuracy of 92% by using a reduction feature based on PCA and KNN for enhanced classification. Based on the results, PC1 accounts for more than 81.7% of the sample variance, and PC2, PC3 and PC4 accounts for 7.7%, 5.87% and 2.39%, respectively. Collectively, PC1, PC2, PC3 and PC4 together has a score of approximately 97.66%, thus the first four principal components represent best variability for overall data set. Through PCA we were able to extract features of our large dataset acquired from the 10 sensor readings. The classification results show that our e-nose using KNN classifier is an effective method for spoilage detection of foods. Thus, the researchers concluded that it is possible to detect spoilage by monitoring the gases emitted by fruits and vegetables inside a refrigerator using Principal Component Analysis (PCA).

V. FUTURE WORKS

To further improve the monitoring of freshness and determining food spoilage inside the refrigerator, the researchers identified the following recommendations: (1) include the study of different food samples since we only focused on banana, grapes, carrots and pechay, (2) Specify which sample is detected as spoiled, (3) Separate the temperature sensor in the array of gas sensors to avoid discrepancy in temperature acquisition due to overheating of the sensing element, or allow the sensors to cool down either by ventilating the system or let it cool down before use, (4) Automate the system to take any corrective action on the food samples and (5) Implement wireless connection in the system for device portability.

VI. REFERENCES

[1] R. Sarno, D. Wijaya (2019). *Recent Development in Electronic Nose Data Processing for Beef Assessment.* Telkomnika, Vol. 17, No.1, pp. 337-348.

[2] R. Mohamed, R. Yaacob, M. Mohamed, T. Dir and F. Rahim (2018). *Food Freshness Using Electronic Nose and Its Classification Method: A Review.* International Journal of Engineering & Technology, Vol. 7, No. 3.28, pp. 49-53

[3] L. Capelli, S. Sironi and R. Del Rosso (2014). *Electronic Noses for Environmental Monitoring Applications.* Sensors, 14, 19979-20007; doi:10.3390/s141119979

[4] G. Green et al (2009). *Monitoring of Food Spoilage with Electronic-Nose: Potential Applications for Smart Homes*, 31st Annual International Conference of the IEEE-EMBS

[5] K. Timsorn, C. Wongchoosuk, P. Wattuya, S. Promdaen and S. Sittichat (2014). *Discrimination of Chicken Freshness using Electronic Nose combined with PCA and ANN.* 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)

[6] A. Amari, N. El Bari and B. Bouchikhi (2007). *Electronic Nose for Anchovy Freshness Monitoring Based on Sensor Array and Pattern Recognition Methods: Principal Component Analysis, Linear Discriminant Analysis and SupportVectorMachine.* Computing, Vol. 6, No.3, pp. 61-67.

[7] I. Tazi, N. Isnaini, M. Mutmainnah and A. Ainur (2018). *Principal Component Analysis (PCA) Method for Classification of Beef and*

Pork Aroma Based on Electronic Nose. Indonesian Journal of Halal Research, Vol. 1, No. 1, pp. 5-8.
[8] N. Linsangan, H. Abubo, R. Matanguihan and M. Natividad (2013). Meat Authentication by Carbon Dioxide and Nitrogen Detection Using Electronic Nose System. Mapua University, Manila, Philippines
[9] N. UI Hasan, N. Ejaz, W. Ejaz and H. Kim (2016). *Meat and Fish Freshness Inspection System Based on Odor Sensing.* Sensors, 12, 15542-15557; doi:10.3390/s121115542a
[10] M. Baietto, A. Wilson (2015). *Electronic-Nose Applications for Fruit Identification, Ripeness and Quality Grading.* Sensors (Basel), 15(1):899-931. doi: 10.3390/s150100899.
[11] J. Lim, N. Linsangan and F. Cruz (2014). *Temperature Compensated Electronic Nose for Fruit Ripeness Determination Using Component Correction Principal Component Analysis.* Mapua University, Manila, Philippines
[12] J. Lim, N. Linsangan and F. Cruz (2014). *A Microcontroller Based Electronic Nose for Fruit Ripeness Determination.* Mapua University, Manila, Philippines
[13] J. Villaverde,, D. Arenga and J. Dela Cruz (2017). *Ripeness Classification of Cocoa through Acoustic Sensing and Machine Learning.* 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, pp. 1-4
[14] J. Villaverde, D. Tagohoy (2019). A Fuzzy Logic Approach for the Determination of Cavendish Banana Shelf Life. TENCON 2018-2018 IEEE Region 10 Conference, Manila, pp. 2467-2472
[15] A. Alejandro, J. Gonzales, J. Yap and N. Linsangan (2018). Grading and Sorting of Carabao Mangoes using Probabilistic Neural Network. AIP Conference Proceedings 2045, 020065; doi:10.1063/s1508078
[16] M. Richardson (2009). Principal Component Analysis, 2-22
[17] D. Kumbhar, D. Singh, M. Abbas, M. Alex and A. Yadav (2018). IoT Based Home Security System Using Raspberry Pi 3. International Journal of Innovative Research in Computer and Communication Engineering, Vol. 6, No. 4, pp.1-8.
[18] A. Shahid, J. Choi, A. Rana, H. Kim (2018). Least Squares Neural Network-Based Wireless E-Nose System Using an SNO2 Sensor Array. Sensors (Basel), 18(5): 1446. doi: 10.3390/s18051446.
[19] M. Kabe, S. Deshmukh, T. Mujawar, V. Bachuwar, L. Deshmukh and A. Shaligram (2015). An Electronic Nose with LabVIEW using SNO2 Based Gas Sensors: Application to Test Freshness of the Fruits. International Journal of Scientific & Engineering Research, Vol. 6, No. 4
[20] S. Benedetti, S. Mannino, A. Sabatini and G. Marcazzan (2004). Electronic Nose and Neural Network use for the Classification of Honey. Apidologie, 35(4):397–402; DOI: 10.1051/apido:2004025