Development of Wavelet Transforms to Predict Methane in Chili using The Electronic Nose

1st Shoffi Izza Sabilla
Departement of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya 60111, Indonesia
shoffi.izza@gmail.com

2nd Riyanarto Sarno
Departement of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya 60111, Indonesia
riyanarto@if.its.ac.id

***Abstract*—Chili (*Capsicum annum L.*) became one of the most popular fruits. In Indonesia, there are various types of chili, but the most commonly purchased by the community are cayenne pepper (*Capsicum frutescens*) and green cayenne pepper. One of the gases in chili is methane. When the body is exposed to methane, the oxygen level will decrease and the lungs will lack oxygen that causes dehydration, resulting in a serious impact on the body.**

This study aims to detect the type of chili that has the most methane content by using E-Nose. One of the E-Nose is an MQ-6 sensor. This sensor can detect, analyze and distinguish methane content.There are seven stages of this research that are, proper sensor selection, training data retrieval that will be used as groundtruth data, signal readout from sensor, noise reduction using low pass filtering and wavelet transform, then normalized with Z-Score, compare Support Vector Machine with Linear Discriminant Analysis, and calculate accuracy, precision, sensitivity with confusion matrix 20 times iteration.

This Study obtained an accuracy of 93% with two classes of cayenne pepper and green cayenne pepper using Linear Discriminant Analysis and 92% with Support Vector Machine. To convert the signal frequency to ppm using Artificial Neural Network obtained ppm. The result is value from green cayenne pepper of 2000 to 6000 ppm whereas in cayenne pepper under 2000 ppm. From the results of experiment's chili that contains the most methane is green cayenne pepper.

***Index Terms*—Chili, E-Nose, Sensor, SVM, LDA, ANN**

I. INTRODUCTION

The demand for Fuel Oil (BBM) of the world from year to year is increasing, which causes oil prices to soar. Many innovations have been made as alternatives or substitutes of fuel oil, one of which is the manufacture of biogas [1] . Biogas gives a positive impact on the environment because of the raw material from organic materials and easily obtainable. The biggest gas content to make biogas is methane and carbon dioxide [2]. This gas content can be obtained from cow dung and livestock manure mixture with the rest of the farm. However, in previous studies, chili can be a raw material of biogas because chili contains methane [3]. Methane is the largest gas content to make biogas.

In Indonesia, chili becomes one of the most popular fruits [4], spicy flavor produced from chili is found in the substance capsaicin [5]. The tongue of the Indonesian people is very fond of the spicy taste. There are various types of chili in Indonesia, but the most commonly purchased by the community are cayenne pepper and green cayenne pepper.

Methane in large quantities is very harmful to the body and the environment [6]. In addition to being one of the causes of earth warming, if eaten will have a negative impact on the human body. Human lungs can run normally if the oxygen level in the body reaches 21%. When the body is exposed to methane, the oxygen level will decrease, and the lungs will lack oxygen that causes dehydration, to cause a serious impact on the body [7].

In this research, will be classified to cayenne pepper and green cayenne pepper by distinguishing methane produced. The output of this research is to determine the two types of chili above which most contain methane. Electronic-nose is used to detect the methane gas content (CH4) contained in chili. An electronic nose is a sensor that can work like a human nose. The human nose can detect gas content in the air, as well as E-Nose but this sensor can also detect, analyze and distinguish the gas content [8]. There are several kinds of E-Nose, but in this study using MQ Family is MQ-6.

The value generated from the sensor is a digital value that will be converted the ratio of existing air resistance and base line air resistance (Rs/Ro). To reduce noise from Rs/Ro value generated using wavelet transforms, Fast Fourier Transform. FFT value in inverse or called low pass filtering. That's one of the processes in the wavelet transform of Discrete Wavelet Transform (DWT) to reduce noise on the signal. Tried one by one of the existing Discrete Wavelet Transform types to determine the best DWT type in this study. To improve accuracy, feature extracts are performed using kurtosis, skewness, mean, and std. The result of feature extraction has different values so it needs to be normalized using Z-score. After normalization, classification is done using Support Vector Machine and Linier Discriminant Analysis. To obtain a good classification method, it can be calculated using a confusion matrix on each classification method.

II. LITERATURE REVIEW

A. Gas Methane in Chilli

In the previous research, has been done to analyze the influence of chili in the manufacture of biogas. The study was conducted for 60 days to detect the methane content produced

TABLE I
TYPE OF DESCRETE WAVELET TRANSFORM

Type of Wavelet	*Family Wavelet*
haar	haar
db	db1, db2, db3, db4, db5, db6, db7, db8, db9, db10
sym	sym2, sym3, sym4, sym5, sym6, sym7, sym8
coif	coif1, coif2, coif3, coif4, coif5
bior	bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8
rbior	rbio1.1, rbio1.3, rbio1.5, rbio2.2, rbio2.4, rbio2.6, rbio2.8, rbio3.1, rbio3.3, rbio3.5, rbio3.7, rbio3.9, rbio4.4, rbio5.5, rbio6.8
meyr	meyr
dmey	dmey
gaus	gaus1, gaus2, gaus3, gaus4, gaus5, gaus6, gaus7, gaus8
cmor	cmor1-1.5, cmor1-1, cmor1-0.5, cmor1-1, cmor1-0.5, cmor1-0.1

by chili. On the 52nd-day, methane production began to look significantly. Methane was found in cayenne pepper with a weight of 12 g/l of 38.1%, and in chili pepper weight of 8 gr/l of 38.538% [3]. In this research, will be measured the content of methane present in chili. Types of chili to be investigated in this study are cayenne pepper and green cayenne pepper.

B. Wavelet Transform

Transforms is the process of analyzing non-stationary signals by separating various characteristics at various scales [9]. Wavelet is a grouping of data in the form of sinusoidal with a variety of frequencies. There are two kinds of wavelet transforms that handle the signal is Fast Fourier Transform and Discret Wavelet Transform. Fast Fourier Transform interprets signals with a certain frequency only into a single frequency band. The weakness in the Fast Fourier Transform is that if the measured data is not accurate, then to distinguish the two signals will be more complicated. Flexibility in Fast Fourier Transform handled in the Discrete Wavelet Transform method, this method is considered more sensitive in showing changes in the form of signals over a certain frequency range and can detect trends and similarities of some signals processed [10].

Discrete Wavelet Transform will convert the signal into two signal classification. That is, the high-frequency signal with high time resolution (*high-pass filter*) and low frequency with low time resolution (*low-pass filter*) [11]. The first Discrete Wavelet Transform action is filtering incoming signals by passing signals on the high-pass filter and low-pass filter, then taking half the sample data from the filter output. And the last one decomposes the signal.

This study will try to use the Fast Fourier Transform which results will be invoked for the low-pass filtering process of the Discrete Wavelet Transform method. Type of Discrete Wavelet Transform can be seen in Table I.

C. Distribution Normal

The result of analysis as a Parametic method becomes true if the treated data has a normal distribution. To ascertain whether the analysis is made correct or not, then the first time the data must be ascertained whether the data to be processed is normally distributed or not.

- Kurtosis
  Kurtosis is the degrees of the frequency curve of a data distribution [12]. There are three kinds of kurtosis, positive kurtosis (Leptokurtik), Platikurtik and negative (Mesokurtik). Leptokurtik kurtosis the tangent of the frequency curve is very high because the data distribution is uneven. Platikurtik is a nearly horizontal frequency curve. Negative Kurtosis is the tilting of the frequency curve within normal limits meaning the data is distributed evenly (not high and not flat) so that the tilting of the frequency curve hardly occurs. Equation 1 is the formula of kurtosis, where μ is the mean value.

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

- Skewness
  Skewness is the slope of the frequency curve of a data distribution [13]. There are two kinds of skewness that are positive and negative. Positive skewness, the slope of the frequency curve of a data distribution tends to be much to the left of the median (the value below the median). So the tail of the slope of the frequency curve is more towards the right. Likewise, in negative skewness, the tail of the slope of the frequency curve is more towards the left. Equation 2 is the formula of skewness, where r_t is the value of each observed data, σ is standart deviation of the sample data, μ is the average of the sample data, and T is the number of sample observations.

$$Skewness(S) = \frac{1}{T\sigma^3}\sum_{t=1}^{T}(r_t - \mu)^3 \tag{2}$$

- Mean
  Mean is the process by which the signal frequency will be centered by summing all the data from a sample group, then divided by the number of data samples. Equation 3 is the formula of skewness, where r_2 is the value of each observed data and T is the number of sample observations.

$$Mean(\mu) = \frac{1}{T}\sum_{t=1}^{T} r_2 \tag{3}$$

- Standart Deviation
  A high value σ indicates that the data value is spreading from its middle value μ. If σ is low, then the value gets clustered at its mean value. Equation 4 is the formula of skewness, where r_2 is the value of each observed data, μ is the average of the sample data, and T is the number of sample observations.

$$Variansi(\sigma^2) = \frac{1}{T}\sum_{t=1}^{T}(r_2 - \mu)^2 \tag{4}$$

D. Normalization

Normalization is an important stage where the range of data will be at the same level or equalize the data range so that no data is more inclined to other attributes. In this study after the process of reducing the noise on the signal, and the value of the truth of a method obtained, the value still has a signal frequency that has not been flat then it needs normalization so that the data equally and easy on the next process that is classification. This study uses standardized score normality or Z-score. Equation 5 is the formula of Z-score, where χ is the value of each observed data and μ is the average of observed data, and σ is standart deviation.

$$Z - Score(z) = \frac{(\chi - \mu)}{\sigma} \tag{5}$$

E. Support Vector Machine

Support Vector Machine (SVM) is one of the most advanced machine learning after Neural Network. This classification technique separates two sets of linear data from 12 different classes [14]. This classification technique has a very convincing performance in predicting the class of a new data [15].

If SVM focuses on optimization issues, Linear Discriminant Analysis (LDA) will analyze the data first before doing the classification [16]. Linear Discriminant Analysis distributes all data well. LDA is one of the classification non-linear techniques. This research will try to perform first analysis on the data owned, then will be classified. The results of the Linear Discriminant Analysis technique will be compared with the SVM classification.

$$f_i = \mu_i C^{-1} x_k^T - \frac{1}{2}\mu_i C^{-1} \mu_i^T + ln(p_i) \tag{6}$$

Equation 6 is the formula of Linear Discriminant Analysis, where μ is the average of all data, while μ_i is the average of group i. C_i is the covariance of group i, and calculated group into one covariance matrix value. p is the prior probability vector, and k is the object of a group i.

III. Result and Discussion

A. Collecting Data

This experiment was carried out using electronic-nose with MQFamily and Arduino as a microcontroller. There are several types of sensors in MQFamily that is MQ6, MQ2, MQ3, etc. In previous research, found methane (CH4) in chili. For that sensor selected is a sensor that can detect methane is MQ 6. Sensors installed in the Arduino microcontroller board in accordance with its input PIN, then Arduino will read analog signal that has been converted into digital value. To get Resistance (Rs) on the sensor can be calculated using digital value (ADC). Equation 7 and Equation 8is the formula to get Resistance.

$$Rs = \frac{Vc - VRL}{VRL} xRL \tag{7}$$

$$VRL = \frac{ADC - Vc}{1023} \tag{8}$$

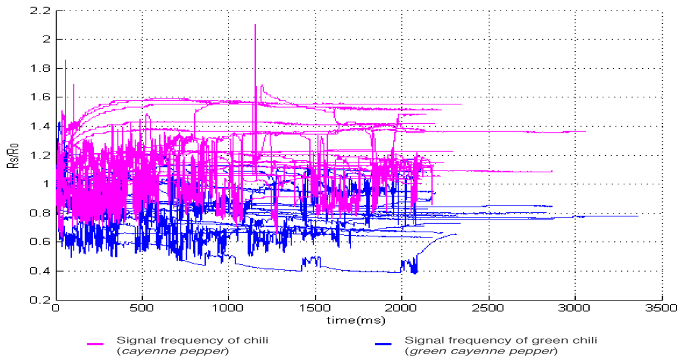

Fig. 1. The Result of resistance value from chili data retrieval on sensor.

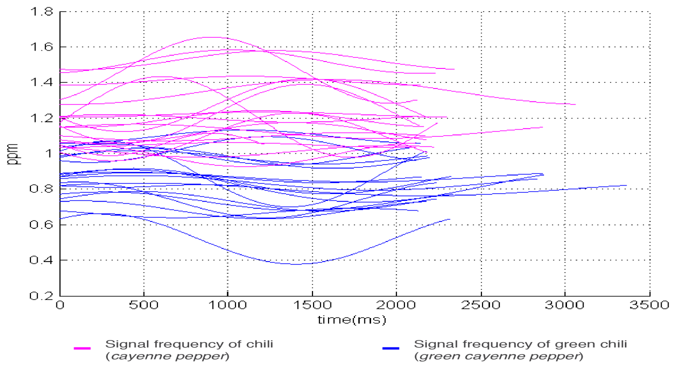

Fig. 2. The result of preprocessing from resistance with SVM

Vc is the voltage generated from the microcontroller board. VRL is the voltage generated by the sensor in the sample space. RL is the resistance value that can be taken from the sensor datasheet with kilohms (Ω). Response value of digital value results from analog to digital conversion (ADC). In Arduino, the microcontroller has a 10-bit resolution where the digital value ranges from 0 to 1023. After the resistance value (Rs) is known, then calculate the Rs/Ro value. Taken data as many as 44 data with details 20 data of cayenne pepper and 20 data from the green cayenne pepper. The experiment was carried out for 35 minutes of each chili. The result of taking Rs/Ro from the sensor can be seen in the Figure 1.

B. Preprocessing

This process is the initial stage in processing input data. Preprocessing transforms unstructured input data into structured data according to need so that it can be analyzed. This stage is very important because the resulting data will improve the data [17]. In a frequency of non-stationary signals, there is noise or noise on the signal so that data quality can decrease. To clear the signal frequency of the interference, can use wavelet transform. Wavelet transforms is a method that can analyze non-stationary signals.

a) Fast Fourier Transform: The workings of Fast Fourier Transform are the high-noise signal frequencies being cut to the closest to the original or average.

b) Low-pass Filtering: There are two types of sensor response values to the changes in gas concentration produced by low-pass and high-pass. In this study using the approximation coefficient of the low-pass filter. The approximation value used depends on the level of decomposition, and the type of wavelet used. The way the low-pass filtering work is by inverse or returning the Fast Fourier Transform value. This research tries to loop the low-pass filter process 10 times to get the best low-pass. The result is the low-pass on the first iteration is good and consistent. The result of low-pass filtering for this study can be seen in Table II. Figure 2 are the result of signal frequency of all data after processing.

TABLE II
THE RESULT OF LOW-PASS FILTERING

Looping	*Best Cut Off*	*Classification*
1	1	SVM
	1	LDA
4	1	SVM
	1	LDA
6	1	SVM
	1	LDA
8	1	SVM
	1	LDA
10	1	SVM
	1	LDA

c) Discrete Wavelet Transform: Having been recognized by the best Low-pass with cut off 1, it is now looking for the best decomposition and wavelet level and the resulting values must be consistent. By the experimenting at the level of decomposition from 1 to 9 with existing wavelet types. And from the experimental results can be seen in Table III, the best are level 7 and level 9 where the type of wavelet used is rbio2.8 and bior3.5.

- There are four types of Threshold selection for de-noising in the wavelet namely Heurure, Rigrsure, Sqtwolog, and Minimaxi. To determine the type of good threshold selection for this research, 100 experiments were conducted and the best to use is the type of threshold Heurure. The result of type of Threshold can be seen in Table III. And this study used Heursure.

TABLE III
THE RESULT OF TYPETHRESHOLD

Wavelet	*Level*	*Type of Threshold*	*Classification*	*Accuracy*
Bior3.5	9	Heursure	SVM	93.552632
Rbio2.8	7		LDA	93.684211
Bior3.5	9	Rigrsure	SVM	92.105263
Rbio2.8	7		LDA	93.157895
Bior3.5	9	Sqtwolog	SVM	93.552632
Rbio2.8	7		LDA	93.684211
Bior3.5	9	Minimaxi	SVM	92.105263
Rbior2.8	7		LDA	93.157895

- After selecting the selection type of threshold, proceed by selecting the threshold, the soft Threshold (s) or hard (h). From the research, it can be seen on Table IV,best to use the selection of soft threshold (s).

TABLE IV
THE RESULT OF THRESHOLD

Type of Threshold	*Threshold*	*Classification*	*Accuracy*
Heursure	S	SVM	93.552632
		LDA	93.684211
	H	SVM	92.105263
		LDA	93.157895

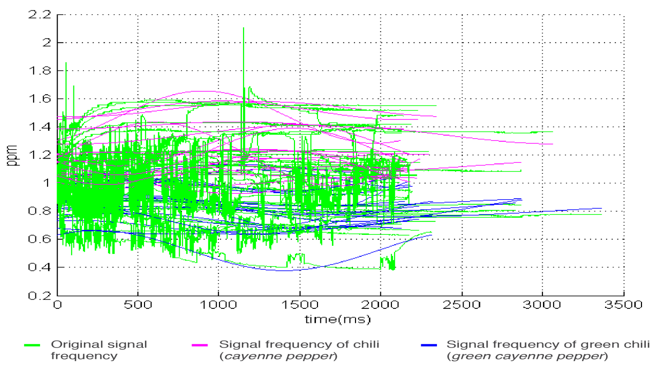

Fig. 3. Comparison between signal frequency after in preprocessing and original signal frequency

- Then, this research tries whether the frequency of the resulting signal requires rescaling. Based on experiments in this study without rescaling (One) will produce better value. The result can be seen in Table V.

TABLE V
THE RESULT OF RESCALING

Threshold	*Rescaling*	*Classification*	*Accuracy*
s	One	SVM	91.710526
		LDA	93.815789
	sln	SVM	86.710526
		LDA	90.657895
	mln	SVM	91.052632
		LDA	91.973684

C. Extraction Feature

The sampling result of 44 chili data gives the value of data distribution level. The Frequency of signals after de-noising still has different or uneven values. In order for the data to not be more inclined to an attribute, the average frequency of the signal using kurtosis, skewness, mean, and standard deviation can be seen in Table VI. However, the resulting value of kurtosis, skewness, mean, and standard deviation has different values or not equal. For that to be normalized so that its value becomes flat, one of the wave signals can be seen in Table VI.

In Figure 3 can be seen the difference from the results of data that has not been in preprocessing and already. The green line is the original data. The magenta colored line is the signal frequency of the cayenne pepper after preprocessing, and the blue line is the signal frequency of green cayenne pepper.

TABLE VI
THE SAMPLE RESULT OF EXTRACTION FEATURE

Distribution	*value*	*Normalization*
Kurtosis	1.4720	-0.2493
Skewness	0.0141	0.5855
Mean	1.4102	1.5599
Standart Deviation	0.0176	-0.8132

D. Classification

The signal frequency is ready for classification, in Table VII showing that the accuracy of the Linear Discriminant Analysis is better than the Support Vector Machine. However, the error rate of SVM is lower than 0.005 LDA.

TABLE VII
THE RESULT OF CLASSIFICATION

Classification	*Accuracy*	*RMSE*
SVM	92.526%	0.031086
LDA	93.947%	0.036191

In Figure 3 can be seen the difference from the results of data that has not been in preprocessing and already. The green line is the original data. The magenta colored line is the signal frequency of the cayenne pepper after preprocessing, and the blue line is the signal frequency of green cayenne pepper.

E. Evaluation

Cross validation is a method used to evaluate data by dividing it randomly into two parts namely, training data and test data. Evaluation is intended to test data mining applications that have been made is good and the percentage of system accuracy increases. The data are tested by a cross-process in which the test data taken at random is made into training data or vice versa. With *k-fold* as much as 10 and looping the cross validation process 100 times to produce consistent accuracy value.

TABLE VIII
EVALUATION OF CLASSIFICATION METHOD

Looping	*Accuracy with SVM*	*Accuracy with LDA*
1	97.368421	94.736842
2	92.105263	93.421053
3	93.8596491	94.736842
5	93.157895	92.631579
10	90.526316	93.421053
15	91.929825	93.859649
20	92.236842	93.947368
25	92.315789	93.157895
50	92.789474	93.052632
100	92.526316	93.263158

Based on Table VIII, shows that cross validation with one repetition of the result is different with the loop counted 20 times. While the repetition results 20 times to 100 times produce a value that is consistent. This means that the result of cross validation generated with a single iteration will change if the experiment returns. For that required cross-validation of 20 times so that the resulting value remains consistent.

F. Confusion Matrix

In order for the method used in this study to correct the signal frequency data in chili, it can use confusion matrix. Confusion matrix is a method used to calculate the accuracy of a method that is being created or that have been implemented. A confusion matrix [18] contains information about actual and predicted classifications done by a classification system. Table IX is the calculation result of a true or false prediction. Can be seen, there are 19 data in the correct prediction and the results are true for cayenne pepper. 17 predicted true data and the result are true for green cayenne pepper.

TABLE IX
THE RESULTS OF THE CALCULATION OF THE TRUE AND FALSE PREDICTIONS

Actual Class	Predicted Class	
	Cayenne Pepper	Green Cayenne Pepper
Cayenne Pepper	19 TP	0 FP
Green Cayenne Pepper	2 FN	17 TN

The output of the confusion matrix is 4,recall, precision, accuracy, and error rate. The accuracy (AC) is the proportion of the total number of predictions that were correct. It is determined using the Equation 9.

$$Accuracy(AC) = \frac{TP + TN}{TP + FN + FP + TN} \quad (9)$$

Precision (P) is the proportion of the predicted positive cases that were correct, as calculated using the Equation 10.

$$Precision(P) = \frac{TN}{FP + TN} \quad (10)$$

The recall is the proportion of positive cases that were correctly identified, as calculated using the Equation 11.

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

TABLE X
THE RESULT OF CONFUSION MATRIX

K	C	A	P	R	RMSE
SVM	Cayenne Pepper	92.526%	0.9202	0.9289	0.0310
	Green Cayenne Pepper	92.526%	0.9289	0.9184	0.0310
LDA	Cayenne Pepper	93.947%	0.9048	1	0.0361
	Green Cayenne Pepper	93.947%	1	0.8847	0.0361

Table X explains that Linear Discriminant Analysis accuracy, precision, and recall rates are higher than Support Vector Machine. K is a classification method. C is a type of chili. A is accuracy. P is precision. R is recall. The RMSE formulas square the error value and divided by the tested data, then giving the roots as a result. Root mean squared error can be shown in Equation 12. Where y_i is the actual data, f_i is the prediction value, and n is the total data.

$$RMSE = \sqrt{\frac{\sum_{i=1} = (y_i - f_i)^2}{n}} \quad (12)$$

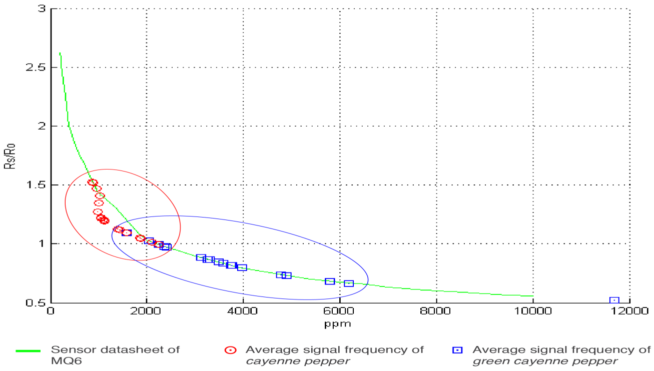

Fig. 4. Convert Rs/Ro to PPM with SVM

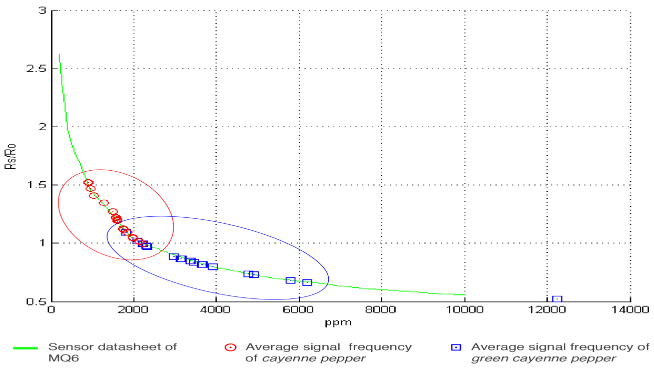

Fig. 5. Convert Rs/Ro to PPM with LDA

G. Convert to PPM

The frequency of the preprocessed signal and the feature extract is converted into methane gas units ie ppm. To change it, this research uses Artificial Neural Network. Artificial Neural Network is a good curve fitting method to convert Rs/Ro value to a certain unit [19]. Using 7 hidden layers and one epoch can produce ppm which can be approximated by Rs/Ro from the MQ 6 sensor datasheet. For classification results using SVM can be seen in Figure 4, and the classification using LDA can be seen in Figure 5 where the green line is datasheet, the red line is cayenne pepper data, and the blue line is green cayenne data.

IV. Conclusion

From the results of research that has been done, it can be concluded that the signal has been in preprocessing, and in extraction can affect the results of the final accuracy. In addition, signals that have been in preprocessing and in extraction help the classification process become easier. The results of these signal frequencies are classified using SVM with 92% accuracy and LDA with 93% accuracy. Based on the results of curve fitting using Neural Network obtained ppm value from green cayenne pepper of 2000 to 6000 ppm whereas in cayenne pepper under 2000 ppm. Thus, it can be concluded that green cayenne pepper contains more methane gas than cayenne pepper.

Acknowledgment

Authors would like to thank Institut Teknologi Sepuluh Nopember and the Ministry of Research, Technology and Higher Education of Indonesia for supporting the research.

References

[1] SOMASHEKAR, R. K.; VERMA, RINKU; NAIK, MANZOOR AHMAD. Potential of Biogas Production from Food Waste in a Uniquely Designed Reactor under Lab Condition. In: The 1st IWWG-ARB Symposium, Hokkaido University, Japan. 2013. p. 18-21.

[2] Ahmad, M. I., Godara, N., Tanweer, S. M. (2015). Implementation and Optimizing Methane Content in Biogas for the Production of Electricity. International Journal of Engineering Research and, V4(06). doi:10.17577/ijertv4is061076.

[3] Cahyaril, K., Sahroni, A. (2015). PENGARUH KONSENTRASI BUAH CABAI MERAH (Capsicum annum L.) DAN BUAH CABAI RAWIT (Capsicum frutescens L.) DALAM PRODUKSI BIOGAS DARI SAMPAH ORGANIK . Jurnal Bahan Alam Terbarukan, 3(1)

[4] Farid, M., N.,Subekti, A. (2012). Review of Production, Consumption, Distribution and Price Dynamics of Chili in Indonesia (2nd ed., Vol. 6, pp. 211-231) (Indonesia, Pusat Kebijakan Perdagangan Dalam Negeri, Badan Pengkajian dan Pengembangan Kebijakan Perdagangan Kementerian Perdagangan-RI).

[5] MATERSKA, Magorzata; PERUCKA, Irena. Antioxidant activity of the main phenolic compounds isolated from hot pepper fruit (Capsicum annuum L.). Journal of Agricultural and Food Chemistry, 2005, 53.5: 1750-1756.

[6] BYARD, Roger W.; WILSON, Gregory W. Death scene gas analysis in suspected methane asphyxia. Am J Forensic Med Pathol, 1992, 13: 69-71.

[7] JO, Jun Yeon, et al. Acute respiratory distress due to methane inhalation. Tuberculosis and respiratory diseases, 2013, 74.3: 120-123.

[8] BAIETTO, Manuela; WILSON, Alphus D. Electronic-nose applications for fruit identification, ripeness and quality grading. Sensors, 2015, 15.1: 899-931.

[9] ANANT, Kanwaldip Singh; DOWLA, Farid U. Wavelet transform methods for phase identification in three-component seismograms. Bulletin of the Seismological Society of America, 1997, 87.6: 1598-1612

[10] AKULOV, L. G. Wavelet filtering in digital signal processing systems. In: Actual Problems of Electron Devices Engineering (APEDE), 2014 International Conference on. IEEE, 2014. p. 5-6.

[11] RIOUL, Olivier; VETTERLI, Martin. Wavelets and signal processing. IEEE signal processing magazine, 1991, 8.4: 14-38.

[12] MANSOUR, Ali; JUTTEN, Christian. What should we say about the kurtosis?. IEEE Signal Processing Letters, 1999, 6.12: 321-322.

[13] ANKARALI, Handan, et al. A bootstrap confidence interval for skewness and kurtosis and properties of t-test in small samples from normal distribution. Balkan Medical Journal, 2009, 2009.4.

[14] SUMATHI, Sai; SIVANANDAM, S. N. Introduction to data mining and its applications. Springer, 2006

[15] XIN, Z. H. O. U.; YING, W. U.; BIN, Y. A. N. G. Signal classification method based on support vector machine and high-order cumulants. Wireless Sensor Network, 2010, 2.01: 48.

[16] XIONG, Tao; CHERKASSKY, Vladimir. A combined SVM and LDA approach for classification. In: Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on. IEEE, 2005. p. 1455-1459.

[17] KIM, Hyuntae, et al. Electronic-nose for detecting environmental pollutants: Signal processing and analog front-end design. Analog Integrated Circuits and Signal Processing, 2012, 70.1: 15-32.

[18] KOHAVI, R. PROVOST, F. Confusion matrix. Machine learning, 1998, 30.2-3: 271-274.

[19] Sabilla, S. I., Sarno, R. (2017, November). Estimating Gas Concentration using Artifical Neural Network for Electronic Nose [Scholarly project]. Retrieved September 8, 2017.