Jurnal Teknologi

Full Paper

A PRELIMINARY STUDY ON DETECTION OF LUNG CANCER CELLS BASED ON VOLATILE ORGANIC COMPOUNDS SENSING USING ELECTRONIC NOSE

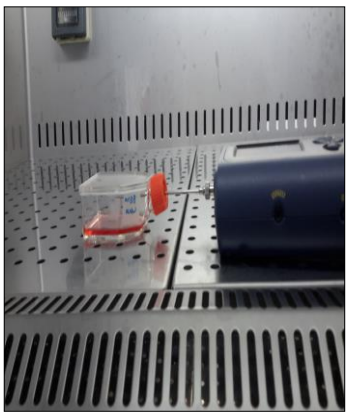
Reena Thriumani[a*], Amanina Iymia Jeffree[a], Ammar Zakaria[a], Yumi Zuhanis Has-Yun Hasyim[b], Khaled Mohamed Helmy[c], Mohammad Iqbal Omar[a], Abdul Hamid Adom[a], Ali Yeon Shakaff[a], Latifah Munirah Kamarudin[a]

[a]Center for Excellence for Advanced Sensor Technology (CEASTech), University Malaysia Perlis
[b]Cell and Tissue Engineering Lab, Department of Biotechnology Engineering, Kulliyah of Engineering, International Islamic University Malaysia (IIUM)
[c]Hospital Tuanku Fauziah (HTF), Jalan Kolam, 01000, Kangar, Perlis, Malaysia.

Article history
Received
19 June 2015
Received in revised form
26 June 2015
Accepted
10 July 2015

*thriumanireena@yahoo.com

Graphical abstract



Abstract

This paper proposes a preliminary investigation on the volatile production patterns generated from three sets of in-vitro cancerous cell samples of headspace that contains volatile organic compounds using the electronic nose system. A commercialized electronic nose consisting of 32 conducting polymer sensors (Cyranose 320) is used to analyze the three classes of signals which are lung cancer cells grown in media, breast cancer cells grown in media and the blank media (without cells). Neural Network (PNN) based classification technique is applied to investigate the performance of an electronic nose (E-nose) system for cancerous lung cell classification.

Keywords: E-Nose, Volatile Organic Compounds (VOCS), in-vitro, lung cancer cells, Probabilistic Neural Network (PNN)

Abstrak

Kertas in membincangkan satu penyiasatan awal keatas corak yang dijana oleh ruapan sebation organic (*VOC*) daripada tiga set 'in-vitro' sel kanser sampel. '*VOC*' daripada '*headspace*' sampel dikumpulkan dengan menggunakan sistem hidung elektronik (E-Nose). E-Nose komersial yang terdiri daripada 32 sensor polimer (Cyranose 320) digunakan untuk menganalisis tiga jenis kelas iaitu sel-sel kanser paru-paru di dalam media, sel-sel kanser payudara di dalam media dan media kosong (tanpa sel-sel). Teknik pengelasan berasaskan 'Neural Network' (PNN) digunakan untuk menyiasat prestasi sistem E-Nose untuk membezakan kanser sel paru-paru.

Kata kunci: E-Nose, Ruapan Sebation Organic (VOCs), In-vitro, kanser sel paru-paru, Probabilistic Neural Network

1.0 INTRODUCTION

World Health Organization's Globocan 2012 estimated worldwide that new cancer cases and cancer related deaths increased to 14.1 million and 8.2 million respectively in 2012, compared to the year 2008 where the number of cancer cases reported as 12.7 million and cancer related deaths were 7.6 million. Due to global population growth and ageing, the number of new cases and cancer related death are predicted to be increased to 19.3 million per year by 2015. The most commonly diagnosed cancers worldwide in the year 2012 was the lung cancer with percentage of 13.0% of the total cases followed by breast cancer (11.9%) and colorectal cancer (9.7%) [1]. Meanwhile, based on information about burden of cancer in ASEAN countries, lung cancer was reported as the most diagnosed and leading cancer related deaths with 98,143 cases and 85,772 deaths in the year 2008 [2].

The common clinical screening methods of lung cancer are normally invasive, apply ionizing radiation and not suited for widespread population screening [3]. Therefore, a new simple, non-invasive, accurate and rapid diagnostic test is needed to replace the conventional methods with comparable sensitivity and specificity for early detection of lung cancer [4].

It is known that odour which contains volatile organic compounds, can be used to diagnose diseases and this concept has been implemented by the Greek and Chinese since 2000 BC [5]. The principle behind this fact is based on cell biology. As a cancer grows, the cell is accompanied by changes in gene or protein expression which leads to peroxidation of cell membrane and emission of chemicals or VOCs from the membrane surface, which can be detected from the headspace of the cell [6][7]. The VOCs can also be monitored via exhaled air breath of patients due to the exchange of cancer related blood chemistry through the lung which can be measured or analyzed [6].

Many researchers since decades back have been focusing on analysis VOCs of exhaled breath of lung cancer patients using VOC analyzer such as gas chromatography-mass spectrometry (GC-MS), solid phase micro-extraction (SPME), proton transfer reaction mass spectrometry (PTR-MS) and selective ion flow tube mass spectrometry [8]. These tools have been very valuable in assisting the researchers to identify the possible biomarkers and understand the biochemistry of human diseases. However, these tools do not appear as routine instrument for clinical diagnosis due to some limitations such as high operation costs, time consuming (for sample preparation) and high need of expertise for effective operation and reliable interpretation of data [9][10]. In that case, the electronic nose, which consists of an array of sensor, was developed and it can fulfill the requirement of clinical applications to accurately detect lung cancer or many other diseases.

The electronic nose system was developed by mimic the human olfactory system using an array of chemical sensors that produce electrical signals and combined with a pattern recognition system. The adsorption of VOCs on the sensor surface leads to changes in the physical properties of sensors such as conductivity, resistance or frequency. This specific change (signals) can be recorded and transformed into digital data which then can be computed based on chemometric analysis [3][11]. Considering that, in this preliminary study, we have investigated the ability of Cyranose 320 to detect lung cancer using lung cancer cultured cells. The A549 (lung cancer cell line) and MCF7 (breast cancer cell line) were cultured and the VOCs released from these cells were sniffed using Cyranose 320 (E-Nose). The VOCs patterns were analyzed using the PNN algorithm to investigate the effectiveness of Cyranose 320 in distinguishing the lung cancer cells from control samples.

2.0 EXPERIMENTAL

There are three sub-sections under this chapter which are materials, instrumentation and chemometric analysis. Figure 1 shows the summary of the methodology used in this study to classify the lung cancer based on VOCs captured using Cyranose320.

2.1 Materials

A549 (lung cancer cell) and MCF7 (breast cancer cell) were provided by Cell and Tissue Engineering Lab, International Islamic University Malaysia. A549 and MCF7 were cultivated in 75cm2 T-flask using complete medium (90% v/v) Dulbecco's Modified Eagle's Medium (DMEM) with (10% v/v) bovine foetal serum (FBS) at 37°C in a humidified atmosphere of 5.0% CO2. The cells were monitored they reach 70%-80% confluence. In sample preparation, for cell seeding process, the 70%-80% confluent adherent cells were harvested using accutase with phosphate buffer solution (PBS) used as washing solution.

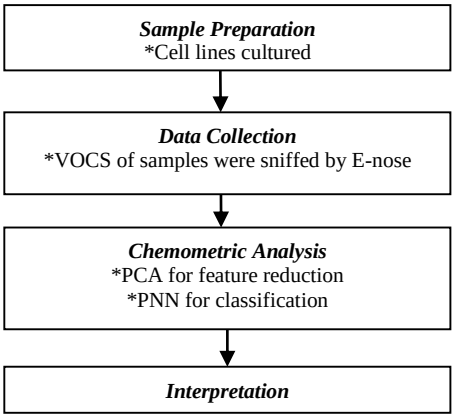


Figure 1 Flowchart of the summary of procedures conducted to classify distinguish the lung cancer based on VOCs

A standardized concentration of each cell line (1 x 10^5 cells/ ml) was seeded into 25cm2 surface treated T-flasks for the E-nose experiment. The samples were prepared in triplicates and incubated for 72 hours in a humidified atmosphere of 5% CO2 at 37°C. The blank medium (without cells) samples were also prepared in triplicates and incubated together with A549 and MCF7 samples as a control data.

2.2 Instrumentation

In this study, the commercially available E-nose, known as Cyrano Science' Cyranose 320 was used. This portable and handheld E-nose contain an array of 32 individual polymer sensors blended with carbon black composite with additional pumps and valves to draw sample of the sensing array. As shown in figure 2, when the sensors interact with the volatile compounds, they swell and lead to changes of the carbon pathway's conductivity and cause the resistance value to increase, which can be monitored as the sensor signal and used to characterize a specific smell. The Cyranose 320 settings for both determination methods and raw data and class training can be stored for additional analysis in Windows based PC using PCnose software [11][12].

2.3 Sampling Procedure

The headspaces of VOCs of the cancerous cell and blank medium in the 25cm2 T-flask were sniffed with Cyranose 320. The snout of Cyranose 320 was inserted into the closed T-25cm2 flasks for 15 minutes. Each measurement was repeated for 3 times. To inspect the VOCs emitted by the cancerous cells at different confluence states, the sampling process was conducted once every 24th hour (1st day) until 72th hour (3rd day).

2.4 Chemometric Analysis

Commercially available softwares; Statistical Package for Social Science (SPSS 17) and Matlab (Matwork, Natrick, MA) were used to process the raw data and identify the effectiveness to improve the effectiveness of E-nose system performance in detecting and identifying the lung cancer cell. Table 1 show the parameters used in the data analysis and classification.

a) Dimensionality Reduction

The 10 end points of the steady state value (Rss) that can be expressed as the stationary values reached during the measurement were extracted from the each response. The high dimensionality of features in the raw data obtained can cause an increase in time and space required for data processing. Therefore, the unsupervised Principal Component Analysis (PCA) techniques were used to improve the accuracy of classification by reducing its dimension while retaining the important features in different basis [13].

Based on Kaiser's rule, the principal components with eigenvalues greater than 1.0 are selected and used as input of the classifier in this study. Table 1 shows the number of principal component selected for each data [14].

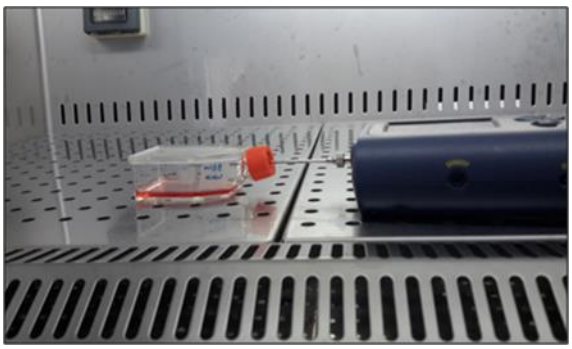


Figure 2 Snout of Cyranose 320 Inserted Into T-25cm2 Flask Containing Cell Culture

b) Classification

In this work, a well applied algorithm in pattern recognition, probabilistic neural network (PNN) classifier is used to classify lung cancer from control samples. PNN neural network algorithm was developed by Specht in 1989 with a simple construction and wide application [15]. PNN which defined as an implementation of Kernel discriminant analysis contains operations which are organized into multi-layered feed forward network with four layers which are input layer, hidden layer, summation layer and output layer [16] as shown in figure 3. Although PNN algorithm required large memory for training, it is widely used as classifier in many applications since it required less training time, able to learn quickly as compared to other than several neural networks model and able to meet to the optimal classifier as the size of the representative training set increases and training samples can be added or removed without extensive retraining [16][17]. Thus, PNN is described and considered as a capable algorithm to be used as lung cancer classifier in this paper.

The spread parameter of PNN algorithm, σ, is crucial to obtain accurate classification. This is because, σ of too small value lead to a solution that does not generalize from the input vectors while σ of too large value can result in large output value of radial basis neurons for all inputs used in the design of the network. In that case, determinations of σ value with the range of 0.1 to 0.9 were conducted for each data set to obtain an appropriate σ with respect to classification accuracy [18]. Table 1 shows the number of principal component selected for each data.

The 10 end points of the steady state value (Rss) that can be expressed as the stationary values reached during the measurement were extracted from the each response. The high dimensionality of features in the raw data obtained can cause an increase in time and space required for data processing. Therefore,

the unsupervised Principal Component Analysis (PCA) techniques were used to improve the accuracy of classification by reducing its dimension while retaining the important features in different basis [13].

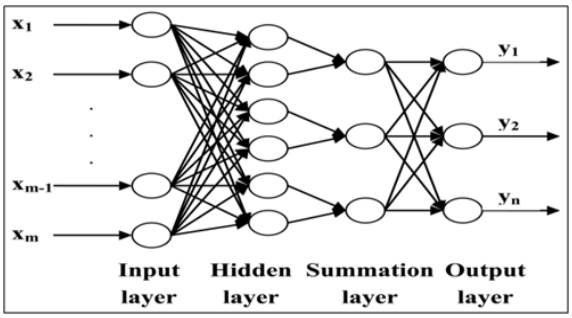


Figure 3 The diagram of PNN built with 4 layers

3.0 RESULTS AND DISCUSSION

In this preliminary study, we used Cryanose 320 E-nose to detect lung cancer cells from breast cancer cells and blank medium with the aid of PNN classifier. The data collected once every 24 hours for 3 days and grouped into 3 classes which are A549 (lung cancer cell line), MCF7 (breast cancer cell line) and blank medium. The PCA was used to reduce the dimensionality of data and speed up the evaluation before PNN classification was conducted. All of the experimental results were averaged over several runs of randomly generated 60/40, training/test splits of the data. The spread factor of PNN, which range starting from 0.1 to 0.9 is conducted on each 4 dataset.

Table 2 demonstrated the result of accuracy, sensitivity and specificity of a PNN classifier on detecting VOCs of lung cancer cell from breast cancer cell and blank medium. All the dataset are checked through spread factor and the results show that outputs successfully rated over 99% of all spread values. However, for spreading factor values 0.7, the output matched highest percentage for all 4 datasets. The percentage of accuracy, sensitivity and specificity on day 1, day 2 and day 3 matched 100% respectively when used spread value 0.7.

Table 1 Properties and parameters of dataset for PNN algorithm

Parameters	Dataset			
	Day 1	Day 2	Day 3	3 days
Total Number of Features	32	32	32	32
Total Samples	450	450	450	1350
After PCA (Features)	5	4	4	5
Training Samples	5x270	4x270	4x270	5x810
Testing Samples	5x180	4x180	4x180	5x540
Classes	3	3	3	3

The accuracy of data from a combination of all 3 days shows 99.98 ± 0.15%, while the sensitivity and specificity records 100% and 99.98 ± 0.23% respectively. For spread value 0.3, the percentage of accuracy, sensitivity and specificity were similar with the outputs from spread value, 0.7; however the percentage of specificity of data on a combination of all data were lesser 0.01% than spread value 0.7.

In nutshell, it can be said that the testing of lung cancer cell detection using an electronic nose system with the aid of PNN classifier was done successfully. These studies show that this method able to rapidly and effectively discriminates the cancer cell types and also from the control which is served as baseline response. The PNN classifier with spread value 0.3 and 0.7 can be used for the best performance as it is proved to be a fast learning algorithm and does not need an iterative training process.

Table 2 The accuracy, sensitivity and specificity of PNN classification

σ Value	Day	Accuracy	Sensitivity	Specificity
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.1	*3*	99.89±0.32	100 ± 0	99.15±0.48
	3 days	99.93 ± 0.1	100 ± 0	99.89±0.15
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.2	*3*	99.86 ± 0.57	100 ± 0	99.73± 0.86
	3 days	99.93±0.1562	100 ± 0	99.89± 0.23
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.3	*3*	100 ± 0	100 ± 0	100 ± 0
	3 days	99.98 ± 0.19	100 ± 0	99.9 ± 0.29
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.4	*3*	99.33±0.57	100 ± 0	98.99±0.87
	3 days	99.98 ± 0.17	100 ± 0	99. 91±0.26
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.5	*3*	99.89±0.35	100 ± 0	99.84±0.49
	3 days	99.97±0.19	100 ± 0	99.96±0.25
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.6	*3*	100 ± 0	100 ± 0	100 ± 0
	3 days	99.96±0.12	100 ± 0	99.94±0.18
	1	100 ± 0	100 ± 0	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.7	*3*	100 ± 0	100 ± 0	100 ± 0
	3 days	99.98±0.15	100 ± 0	99.98±0.23
	1	99.78±0.29	99.84±0.87	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.8	*3*	100 ± 0	100 ± 0	100 ± 0
	3 days	99.88±0.12	99.78±0.29	99.94±0.18
	1	99.86 ± 0.29	99.54± 0.87	100 ± 0
	2	100 ± 0	100 ± 0	100 ± 0
0.9	*3*	100 ± 0	100 ± 0	100 ± 0
	3 days	99.81 ± 0.17	99.67± 0.28	99.89± 0.23

Considering that, the spread value, σ, 0.3 and 0.7 can be used for all the datasets to achieve efficient lung cancer classification system. The PNN analysis applied in this study reached 100% classification and these phenomena might occur because of the dataset was relatively small and the variability associated with the classification is still high. Therefore, for more accurate quantification of the general performance of PNN using e-nose data, further experimentation with larger datasets is necessary [19].

4.0 CONCLUSION

This study proposed a lung cancer cell VOCs detection system which consists of a commercially available electronic nose (Cyranose 320) and pattern recognition algorithm (using PNN). The results proved that this proposed system is able to detect lung cancer VOCs by achieving over 99% of classification accuracy. Although the current pattern classification technique produces pleasing results, but further study may improve the classification accuracy and speed of the operations by using different types of pre-processing and classification methods. Future work will be extended to a larger number of samples where the normal lung cancer cell lines and other types of lung cancer cell lines beside A549 will be investigated.

Acknowledgement

This work is supported by Exploratory Research Grant Scheme by Ministry of Education Malaysia (ERGS 9010-00018). The author would like to thank the Animal Cell and Tissue Culture Lab, Department of Biotechnology Engineering, International Islamic University Malaysia and Hospital Tuanku Fauziah, Kangar, Perlis for the great collaboration.

References

[1] Ferlay.J, Soerjomataram.J, Ervik.M, Dikshit.R, Eser.S, Mathers.C, Rebelo.M, Parkin.DM, Forman.D, and Bray.F. 2013. Cancer Incidence and Mortality Worldwide: IARC Cancer Base GLOBOCAN 2012 Cancer Incidence and Mortality Worldwide: IARC Cancer Base. (11)11.

[2] M. Kimman, R. Norman, S. Jan, D. Kingston, and M. Woodward. 2012. The Burden of Cancer in Member Countries of the Association of Southeast Asian Nations (ASEAN). *Asian Pacific J. Cancer Prev*. 13(2): 411–420.

[3] C. Brunner, W. Szymczak, V. Höllriegl, S. Mörtl, H. Oelmez, a Bergner, R. M. Huber, C.Hoeschen, and U.Oeh. 2014. Discrimination Of Cancerous And Non-Cancerous Cell Lines By Headspace-Analysis With PTR-MS. *Anal. Bioanal. Chem.* (397)6: 2315–2324.

[4] U. Kalluri, M. Naiker, and M. A Myers. 2014. Cell Culture Metabolomics In The Diagnosis Of Lung Cancer-The Influence Of Cell Culture Conditions. *J. Breath Res.* (8)2: 1-10.

[5] R. Fend, A. H. J. Kolk, C. Bessant, P. Buijtels, P. R. Klatser, and A. C. Woodman, 2006. Prospects For Clinical Application Of Electronic-Nose Technology To Early Detection Of Mycobacterium Tuberculosis In Culture And Sputum. *J. Clin. Microbiol*. (44) 6: 2039–2045.

[6] G. Peng, M. Hakim, Y. Y. Broza, S. Billan, R. Abdah-Bortnyak, a Kuten, U. Tisch, and H. Haick, 2010. Detection Of Lung, Breast, Colorectal, And Prostate Cancers From Exhaled Breath Using A Single Array Of Nanosensors. Br. *J. Cancer* (103)4: 542–551.

[7] D. A. P. Daniel, K. Thangavel, and R. S. C. Boss. 2012. A Review Of Early Detection Of Cancers Using Breath Analysis. *Int. Conf. Pattern Recognition, Informatics Med. Eng. Mar*. 433–438.

[8] R. Thriumani, A. Zakaria, M. I. Omar, A. H. Adom, A. Y. M. Sharaff, L. M. Kamaruddin, N. Yusuf, and K. M. Helmy. 2014. Review On Exhaled Volatile Organic Compounds From Lung Cancer And Advances Of E-Nose As Non-Invasive Detection Method. *IEEE 5th Control and System Graduate Research Colloquium*. 213–218.

[9] Manolis.A. 1983. The Diagnostic Potential Of Breath Analysis. Clin. *Chem*. 29: 5–15.

[10] A. D. Wilson and M. Baietto. 2011. *Advances In Electronic-Nose Technologies Developed For Biomedical Applications Sensors (Basel.)* (11)1: 1105–1176.

[11] T. Mohamed, A. Sabeel, S. E. Eluwa, and S. Mohamed. 2000. *Detection of Volatile Compounds in Urine*. 3–6.

[12] R. Dutta, E. L. Hines, J. W. Gardner, and P. Boilot. 2002. Bacteria Classification Using Cyranose 320 Electronic Nose. *Biomed. Eng. Online*. 1: 4.

[13] R. Palaniappan and K. V. R. Ravi. 2006. Improving Visual Evoked Potential Feature Classification For Person Recognition Using PCA And Normalization. Pattern Recognit. *Lett*. 27(7): 726–733.

[14] Jolliffe.I.T. 2002. *Principal Component Analysis*. Springer Series in Statistics.

[15] C. Ji. D, F. Song. B, and H. Yi. H. 2009. Equipment Fault Diagnosis based on Probabilistic Neural Networks and Simulation Analysis. Fire Control Command Control. 34: 82–85.

[16] M. Mishra, A. R. Jena, and R. Das, 2013. A Probabilistic Neural Network Approach for Classification of Vehicle. 2(7): 367–371.

[17] N. Bhattacharyya and A. Jana 2008. *Incremental PNN Classifier for a Versatile Electronic Nose* 242–247.

[18] R. Antony, M. S. G. Nandagopal, S. Rangabhashiyam, and N. Selvaraju. 2014. Probabilistic Neural Network Prediction Of Liquid- Liquid Two Phase Flows In A Circular Microchannel. *J. Sci. Ind. Res*. 73: 525–529.

[19] H. Zhang, M. Ö. Balaban, and J. C. Principe. 2003. Improving Pattern Recognition Of Electronic Nose Data With Time-Delay Neural Networks. *Sensors Actuators B Chem*. 96(1–2): 385–389.