Rapid Identification Method of Aerobic Bacteria in Diabetic Foot Ulcers Using Electronic Nose

Article *in* Journal of Computational and Theoretical Nanoscience · January 2014

DOI: 10.1166/asl.2014.5306

CITATIONS
16

READS
382

12 authors, including:

Azian Azamimi Abdullah
Universiti Malaysia Perlis
95 PUBLICATIONS **812** CITATIONS
SEE PROFILE

Nurlisa Yusuf
Universiti Malaysia Perlis
21 PUBLICATIONS **165** CITATIONS
SEE PROFILE

Ammar Zakaria
Universiti Malaysia Perlis
283 PUBLICATIONS **2,790** CITATIONS
SEE PROFILE

Ye Htut
Mercy Malaysia
71 PUBLICATIONS **1,066** CITATIONS
SEE PROFILE

Advanced Science Letters
Vol. 1, 37–41, 2014

Rapid Identification Method of Aerobic Bacteria in Diabetic Foot Ulcers using Electronic Nose

Azian Azamimi Abdullah[1], Tan Woei Jing[1], Chua Ai Sie[1], Nurlisa Yusuf[1], Ammar Zakaria[1], Mohammad Iqbal Omar[1], Ali Yeon Md Shakaff[1], Abdul Hamid Adom[1], Latifah Munirah Kamarudin[1], Yeap Ewe Juan[2], Amizah Othman[2] and Mohd Sadek Yassin[2]

[1] Center of Excellent for Advanced Sensor Technology (CEASTech),
Universiti Malaysia Perlis (UniMAP),
02600 Jejawi, Perlis, Malaysia

[2] Hospital Tuanku Fauziah (HTF)
Jalan Kolam, 01000 Kangar, Perlis, Malaysia

Diabetic foot ulcer (DFU) is an infection, ulcer or destruction of deep tissue associated with neurological abnormalities, musculoskeletal deformities and various degrees of peripheral vascular disease of lower limb. In this study, an electronic nose is proposed to recognize types of bacteria in the diabetic foot ulcer on patients with the aid of data analysis using PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis) as the first step for feature extraction. Bacteria classification is also enhanced by pattern analysis using FFBP (Feed Forward Back Propagation) and SOM (Self Organizing Map) of ANN (Artificial Neural Network). Experimental results show that the use of electronic nose in identifying bacteria on diabetic foot ulcers works effectively as the high accuracy obtained for the bacteria classification.

Keywords: Diabetic Foot Ulcer, Electronic Nose, Bacteria Infection, Principle Component Analysis, Linear Discriminant Analysis, Feed Forward Back Propagation, Self Organizing Map, Artificial Neural Network.

1. INTRODUCTION

Diabetic ulcers are the most common cause of no traumatic foot injuries leading to lower extremity amputation in the industrialized world and modern era. The diabetic foot ulcer infections are polymicrobial in natura1. There are several conventional diagnosis methods used to decide the diabetic foot ulcer, such as microbiological sampling, hematological and biochemical markers, radiological diagnosis and clinical diagnosis of infection. The most traditional method is rubbing the wound surface with cotton swab to determine the causative agents of a wound infection[1-3]. Bacterial infections caused by the growth of certain micro-organisms known as pathogens which tend to afflict our upper respiratory system[3]. Diabetic foot ulcers commonly have these bacteria such as Staphylococcus Aureus, Escherichia Coli, and Pseudomonas Aeruginosa.

*Email Address: azamimi@unimap.edu.my

Although some of these bacterial genera were found to be common between intact skin and wounds, however others organisms that have been isolated have high percentage in the diabetic foot patient. Those bacteria are Klebsiella Pneumoniae, Proteus Mirabilis, Streptococcus Pyogenes, and Enterobacteraciae.

E-nose was proposed to be trained with qualified samples to build a database of reference. E-nose can recognize new sample of bacteria by comparing volatile compounds fingerprint to those contained in its database and perform pattern-recognition analysis rapidly. E-nose have the commercialize potential in the world-wide. By using PCA, LDA, FFBP and SOM, bacteria classifications can be done.

This paper is organized as follows. In Section 2, literature reviews are done on E-nose, PCA, LDA, FFBP and SOM. In Section 3, the methodology of this project is presented. Section 4 presents the experimental results of

1936-6612/2011/4/400/008
doi:10.1166/asl.2011.1261

bacteria classification using both statistical analysis methods (PCA and LDA) and also artificial neural networks (FFBP and SOM).

2. LITERATURE REVIEW

Specific biomedical E-nose applications range from uses in biochemical testing, blood-compatibility evaluations, disease diagnoses, and drug delivery to monitoring of metabolic levels, organ dysfunctions, and patient conditions through telemedicine[4]. Generally, E-Noses include three major parts which is a sample delivery system, a detection system and a computing system. These parts are the sub component of the electronic nose. The main components consist of sensing system and pattern reorganization system. The sample delivery system enables the generation of the head space (volatile compounds) of a sample, which is the fraction analyzed. For the detection system, the most commonly used sensors include MOS (Metal Oxide Semiconductors), CP (Conducting Polymers), quartz crystal microbalance, SAW (Surface Acoustic Wave), and MOSFET (Field Effect Transistors).The next sub component is the computing system. Computing system works to combine the responses of all the sensors, which represents the responses for the data treatment[5]. Figure 1 show the schematic of an E-nose.

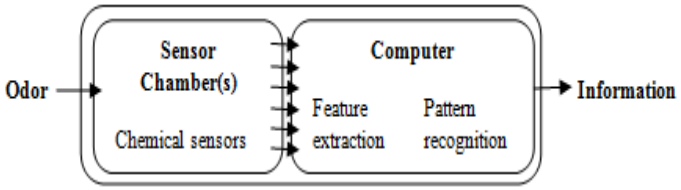

Fig.1. <Schematic of an Electronic Nose>

PCA is techniques that reduce the data into two dimensions by explore correlations between samples or conditions. The correlations are others than the way of identifying patterns in data. It is also expressing the data in such a way as to highlight their similarities and differences[6, 7]. PCA projects the data along the directions where the data varies the most. The goal of PCA is to 'summarize' the data; it is not considered a clustering tool.

LDA is also a commonly used technique for data classification and dimensionality reduction. The aim of LDA is to handle easily the case where within-class frequencies are unequal and their performances have been examined on randomly generated test data[8]. The standard LDA will be seriously degraded if there are only a limited number of observations, N compared to the dimension of the feature space, n[9]. From the literature review done, LDA is more preferred over PCA because LDA deals directly with discrimination between classes, since the PCA deals with the data in its entirety for the principal components analysis without paying any particular attention to the underlying class structure[11].

Feed-forward neural network is a supervised neural network which it knows the desired output and adjusting of weight coefficients is done in such way, that the calculated and desired outputs are as close as possible[10]. In back propagation network, the first step is to calculate the error in the gradient descent and propagates it backwards to each neuron in the output layer, then hidden layer. In the second step, the weights and biases are then recomputed, and the output from the activated neurons is then propagated forward from the hidden layer to the output layer[12].

Self Organizing Map is one of the unsupervised neural network methods for classifying and clustering a set of data. It is transforming the incoming signal pattern from arbitrary dimension into a two-dimensional discrete map, and to perform this transformation adaptively in a topological ordered way such as result represented in U-matrix. SOM is an unsupervised algorithm that works with nonlinear data set. The most advantages of SOM is that it have an excellent ability to visualize high dimensional data of 1 or 2 dimensional space makes it unique especially for dimensionality reduction[13].

3. METHODOLOGY

Figure 2 show the flowchart of the experiment. First, the bacteria are cultured in blood Agar medium at Hospital Tunku Fauziah, Kangar, Perlis with the guide of microbiologist from the hospital. The types of bacteria used in this study are the Klebsiella Pneumoniae, Proteus Mirabilis and Streptococcus Pyogenes. One species for each WILD and STANDARD type bacteria were cultured in 6 samples. WILD type bacteria is the bacteria that directly taken from patient foot ulcers while for STANDARD type bacteria, is the pure bacterium from the ATCC strain. Table 1 shows the total number of bacteria sample. There are 2 batches of bacteria sample for different week. The blank samples which only consist of medium without bacteria used as a control for this experiment.

The equipment used for culture the bacteria is probe and incinerator. The inoculators loop must be heated first by using incinerator to make sure it is clean and sterile from other microorganism and let it cool. Total of 3 strike phases were done on each plate using zigzag pattern during culturing. Next, the mono-culturing bacteria for each species of bacteria will be left for 6 hours at the temperature of 37°C in the incubator to observe the bacteria growth. The growth bacteria samples that had been cultured are tested using E-nose to collect data each consecutive of 6 hours until 36 hours for Batch 1 and Batch 2. After data collection, all data will be further analyzed using statistical analysis methods which are Principal Component Analysis and Linear Discriminant Analysis. Then, artificial neural network will be implemented for the classification purpose.

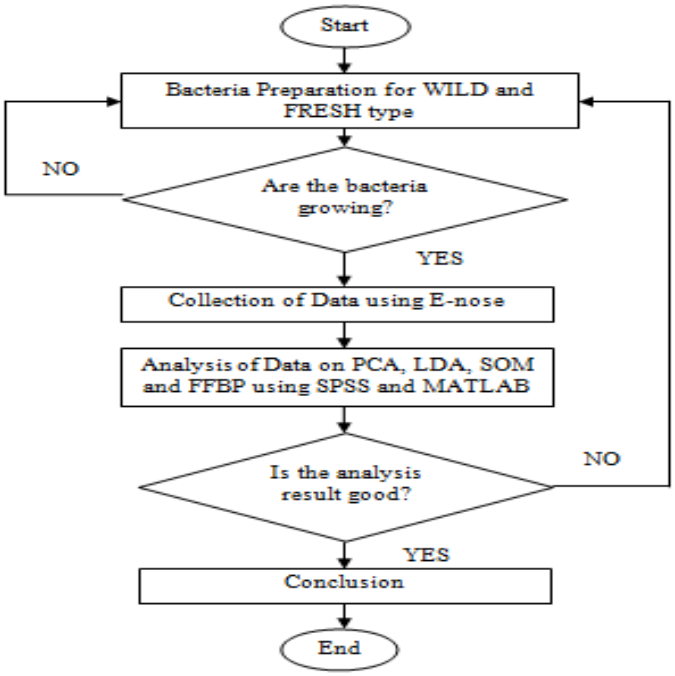

Fig.2. <Flowchart of the Experiment>

Table.1. Number of Bacteria Sample

Types / Bacteria	WILD	STANDARD
K. pneumonia	6	6
P. mirabilis	6	6
S. pyogenes	6	6
BLANK	1	1

4 EXPERIMENTAL RESULTS

The experiment was done for 2 batches for WILD and STANDARD type bacteria with data collection of every consecutive of 6 hours. In this study, the results represented here are focused on combination of data collection of Batch 1 and Batch 2 at 6 hours and 24 hours for each WILD type of bacteria. In Fig. 3 and Fig. 4, red colour plot represents the K. pneumoniae, green colour plot represents the P. mirabilis, and S. pyogenes is represented by blue colour plot whereas the pink colour plot is the BLANK. Figure 3 shows the LDA plot have 100% classification which proves that E-nose is able to recognize and classify the bacteria effectively although the S. pyogenes and BLANK are slightly overlapping. This shows that the odour of S. pyogenes and BLANK are almost the same. The Discriminant function for 6 hours in Fig. 3 is 56.5%, 27.5% and 12.6% respectively for discrimination function 1, 2 and 3. In Fig. 4, the LDA plot for 24 hours has 99.7% of classification with the Discriminant function 1 of 74.5%, Discriminant function 2 of 17.1% and 6.5% for Discriminant function 3. Both for 6 hours and 24 hours LDA plot of WILD type, the clustering and classification of bacteria can be seen clearly.

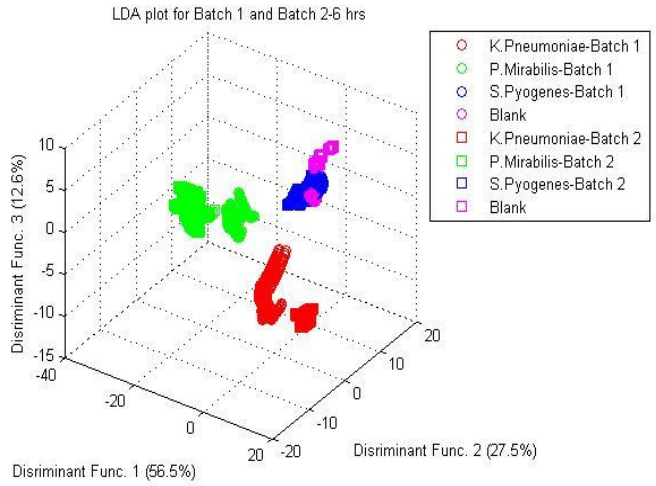

Fig.3. <LDA Plot for WILD Type Bacteria at 6 Hours>

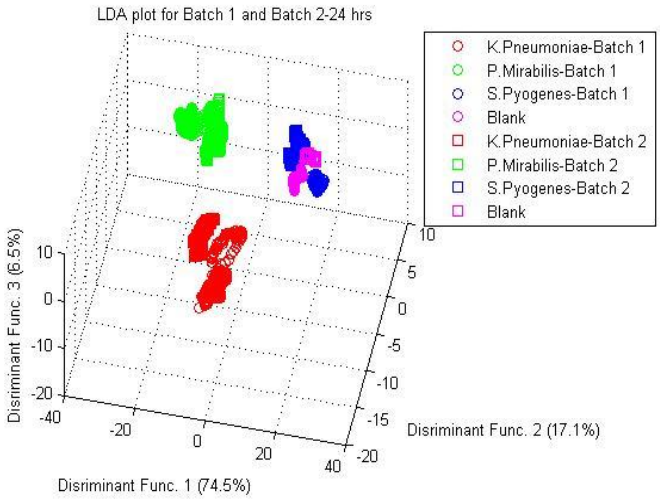

Fig.4. <LDA Plot for WILD Type Bacteria at 24 Hours>

Analysis results of data using PCA are shown in Fig. 5 and Fig. 6 for 6 hours and 24 hours respectively. Red colour plot represents the K. pneumonia, green colour plot represents the P. mirabilis, and S. pyogenes is represented by blue colour plot whereas the black colour plot is the BLANK for Fig. 5 and Fig. 6. For 6 hours, PCA plot have 97.8% of classification with 84.9% principle component 1, 9.7% for principle component 2 and 3.2% for principle component 3. In Fig. 6, 24 hours data has the percentage of classification of 97.1% with the principle component 1, 2 and 3 of 79.8%, 15.5% and 1.8% respectively. Better classification and clustering can be obtained by using U-matrix of SOM compared to LDA plot and PCA plot. In Fig. 7, the bacteria were classified well into their group for batch 1 and batch 2 at 6 hours using LDA data as input to SOM. Figure 8 shows the clustering result using which PCA data as input to SOM for batch 1 and batch 2 at 6 hours. It is observed that the result obtained is not classified properly compared to LDA data as input to SOM as shown in Fig. 7. However, through U-matrix, it performs well classification than PCA plot. Figure 9 and Fig. 10 represent the U-matrix of 24 hours which take LDA data and PCA data as input to SOM respectively.

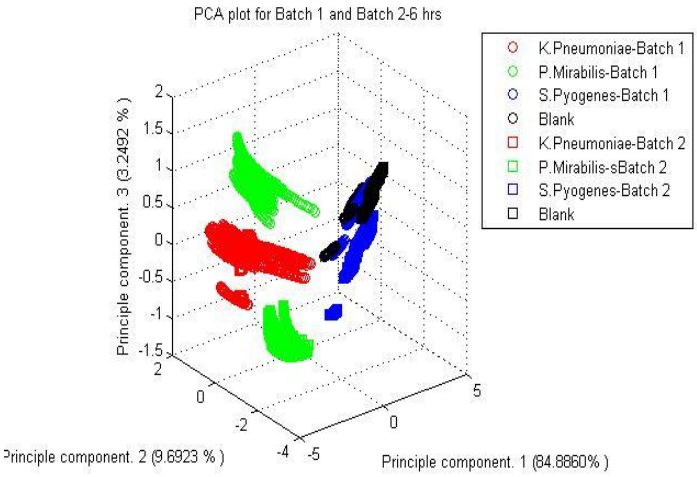

Fig.5. <PCA Plot for WILD Type Bacteria at 6 Hours>

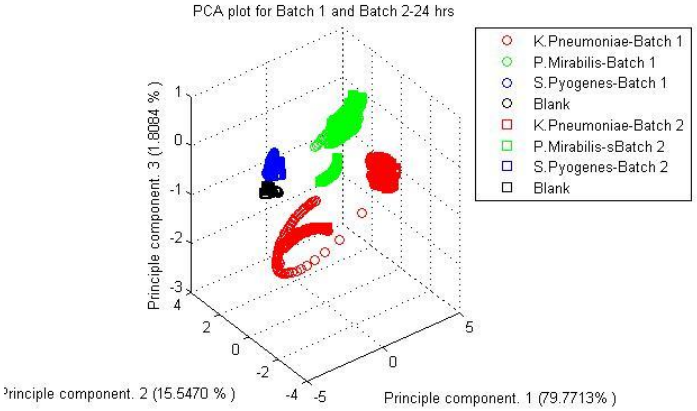

Fig.6. <PCA Plot for WILD Type Bacteria at 24 Hours>

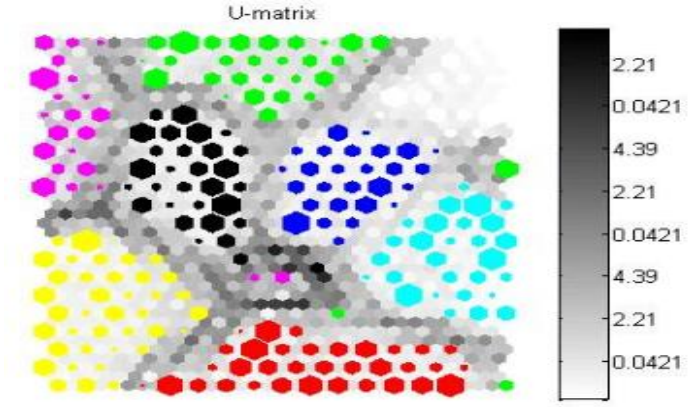

Fig.7. <U-Matrix for LDA WILD Type Bacteria at 6 Hours>

For 24 hours, U-matrix SOM well classifies the bacteria into the group for LDA and PCA data as input data. For SOM, the map size is [20, 17] for U-matrix. SOM data input from LDA and PCA analysis consists of total 4800 samples of three bacteria group and one BLANK. The U-matrix shows distances between neighboring units (the grayscale column) and thus visualizes the cluster structure of the map. High values on the U-matrix mean large distance between neighboring map units, and thus indicate cluster borders. Furthermore, validation of the quality of the SOM can be done by determining the quantization error and topographic error which represented in Table 2. The smaller the quantization error, the smaller the average distance from the vector data to the prototypes, and that means that the data vectors are closer to its prototypes. Both of this error determination evaluates the fitting of the neural map to the data.

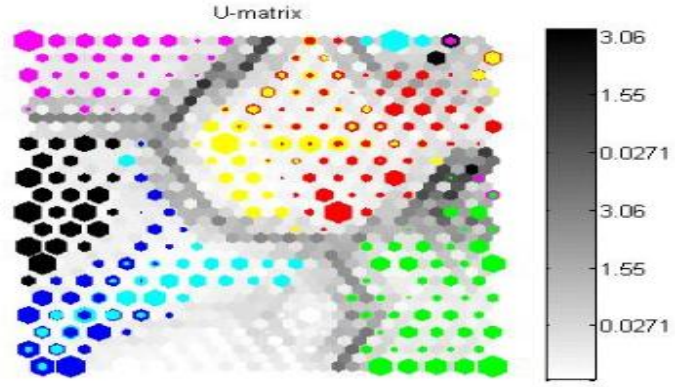

Fig.8. <U-Matrix for PCA WILD Type Bacteria at 6 Hours>

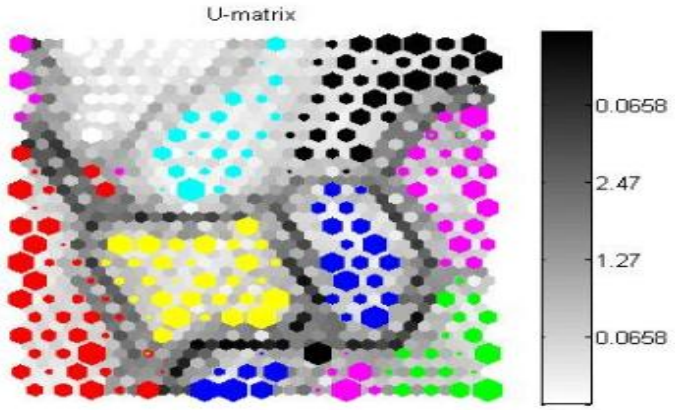

Fig.9. <U-Matrix for LDA WILD Type Bacteria at 24Hours>

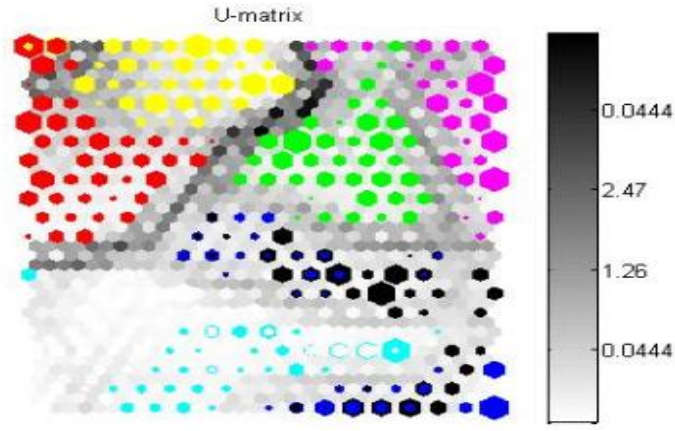

Fig.10. <U-Matrix for PCA WILD Type Bacteria at 24 Hours>

Percentage classification of LDA and PCA can be improved by using FFBP. Taking 25% of the data for the validation set, 25% for the test set and 50% for the training set. Table 3 summarizes the percentage classification of LDA, PCA and FFBP.

Table.2. Experimental Result of Quantization Error and Topographic Error.

Batch 1 and Batch 2		6 Hours	24 Hours
LDA + SOM	Quantization Error	0.4134	0.4921
	Topographic Error	0.0615	0.0719
PCA + SOM	Quantization Error	0.1897	0.2544
	Topographic Error	0.0894	0.0821

Table.3. Experimental Result of Percentage Classification with Different Methods.

Batch 1 and Batch 2	6 Hours	24 Hours
LDA (%)	100	99.7
PCA (%)	97.8	97.1
LDA + FFBP (%)	100	100
PCA + FFBP (%)	98.9	99.75

Table 3 shows the percentage classification of bacteria increased by applying FFBP with the input data from LDA and PCA. For 6 hours, the percentage classification for PCA is 97.8%, however by using FFBP, the percentage classification is increased to 98.9%. For percentage classification of LDA and FFBP at 6 hours, the percentage classification is 100% maximum. As for 24 hours, percentage of classification of LDA which is 99.7% increased to 100% using FFBP. This circumstance also applied to 24 hours using PCA which is 97.1% increased to 99.75% using FFBP. Overall, the percentage of classification can be improved by using supervised method of ANN using FFBP.

5. CONCLUSIONS

From the experimental results, it can be concluded that all bacteria (Klebsiella pneumonia, Proteus mirabilis and Streptococcus pyogenes) are classified successfully. Considering in term of results visually on the plot, LDA plot generates better view than PCA plot. Besides that, the percentage classification for LDA is also better than PCA as for 6 hours, the LDA perform percentage classification of 100% compared to 97.8% of PCA. As for 24 hours, LDA have 99.7% classification more than 97.1% of PCA. Thus, can be conclude that LDA is better than PCA in term of plot visualize and percentage classification. Furthermore, for percentage classification of LDA and PCA that is not reaching the maximum thus it can be improved by the FFBP. Majority of the percentage classification is improved by the FFBP as in 6 hours, percentage classification of PCA were increased from 97.8% to 98.9%. Linear regression between the network outputs and the corresponding targets is used to determine the network response. Linear regression between the network outputs and the corresponding targets is used to determine the FFBP network response. Other than PCA and LDA, U-matrix SOM shows a better visual clustering of bacteria using data reduction from PCA and LDA. This experiment strongly proven that the rapid detection of diabetic foot ulcer can be detected as fast as 6 hours compared to 24 hours. E-nose is able to detect, recognize and classify the bacteria at 6 hours well and with high percentage of classification.

ACKNOWLEDGMENTS

This work is supported by the ERGS grant 9010-00010 and 9010-00015 by Ministry of Higher Education, Malaysia. The authors would like to thanks to the Director General of Health, Malaysia for permission to publish this paper.

REFERENCES

[1] Ozer B., Kalaci A, Semerci E, Duran N, Davul S and Yanat A. N. Infections and aerobic bacterial pathogens in diabetic foot, African Journal of Microbiology, Research, (2010) Vol. 4(20) 2153-2160.
[2] D. T. Williams, J. R. Hilton, and K. G. Harding. Diagnosing Foot Infection in Diabetes, Wound Healing Research Unit, (2004). 3-14.
[3] Alphus D. Wilson, Manuela Baietto. Advances in Electronic-Nose Technologies Developed for Biomedical Applications, vol20 (2011) 4-17.
[4] Fabrizio Davide, Martin Holmberg, Ingemar Lundström. Virtual olfactory interfaces: Electronic noses and olfactory displays, (2001) 1-3.
[5] Simona Benedettia, Saverio Manninoa, Anna Gloria Sabatinib, Gian Luigi Marcazzanb. Electronic nose and neural network use for the classification of honey, (2003) 32-47.
[6] Ying Wu. Principal Component Analysis and Linear Discriminant Analysis, IL 60208 (2008) 5-21.
[7] Lindsay I Smith. A tutorial on Principal Components Analysis, (2002) 2-23.
[8] S. Balakrishnama, A. Ganapathiraju Theory of Linear Discriminant Analysis, (2007) 1-8.
[9] Aamir Khan, Hasan Farooq. Principal Component Analysis-Linear Discriminant Analysis Feature Extractor for Pattern Recognition, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 2, (2011) 1694-0814.
[10] L. Ma, K. Khorasani. A New Strategy For Adaptively Constructing Multilayer Feedforward neural networks, (2002) 3-20.
[11] Aleix M. Martinez, Avinash C. Kak. Pattern Analysis and Machine Intelligence, (2001) Vol. 23, No.2, 228-233.
[12] Neural Networks Toolbox (4.0) User's Guide: The MathWorks, Natick, MA, (2000) 2-2–2-22, 5-2–2-56.
[13] Mr. Vivek A. Patole, Mr. V. K. Pachghare, Dr. Parag Kulkarni. Self Organizing Maps to Build Intrusion Detection System, International Journal of Computer Applications (0975– 8887), (2010), Vol 1 – No. 8, 4.

Received: 15 May 2013. Accepted: 18 August 2013