chemistry proceedings

MDPI

Application of electronic nose technology as a promising non-invasive tool for breath analysis of patients with liver cirrhosis and gastric cancer †

Omar Zaim [1,2], Nezha El Bari [2], Naoual Lagdali [3], Imane Benelbarhdadi [3], Fatima Zohra Ajana [3], and Benachir Bouchikhi [1,*]

[1] Sensor Electronic & Instrumentation Group, Department of Physics, Faculty of Sciences, Moulay Ismaïl University of Meknes, B.P. 11201, Zitoune, Meknes, Morocco; benachir.bouchikhi@gmail.com
[2] Biosensors and Nanotechnology Group, Department of Biology, Faculty of Sciences, Moulay Ismaïl University of Meknes, B.P. 11201, Zitoune, Meknes, Morocco; n_elbari@hotmail.com
[3] Department of Medical Gastroenterology C, Ibn Sina Hospital, Mohammed V University, Rabat, Morocco; ajafatimz@yahoo.fr
* Correspondence: benachir.bouchikhi@gmail.com; Tel.: +212 535 53 88 70; Fax: +212 535 53 68 08; Postal address: B.P. 11201, Zitoune 50003 Meknes
† CSAC2021: 1st International Electronic Conference on Chemical Sensors and Analytical Chemistry, online, 01 – 15 July 2021.

Abstract: This study examined the ability of an electronic nose (e-nose) based on metal-oxide semiconductor sensors combined with chemometrics methods to discriminate between liver cirrhosis (LCi), gastric cancer (GCa) patients, and healthy controls (HC). For this purpose, principal component analysis (PCA), discriminant function analysis (DFA), and support vector machines (SVM) are used for data processing of multivariable responses generated by the sensor arrays. The results showed good discrimination between the three health states. This study reveals that e-nose technology based on exhaled breath analysis could be an effective non-invasive way to distinguish the three studied groups.

Keywords: Liver cirrhosis; gastric cancer; volatile organic compounds; exhaled breath analysis; electronic nose; pattern recognition methods.

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Chem. Proc.* **2021**, *3*, x. https://doi.org/10.3390/xxxxx

Published: date

1. Introduction

Diseases are a constant threat to human life and well-being, with over 200 distinct types affecting up to 60 human organs [1].

Liver cirrhosis is an extended phase of hepatic fibrosis accompanied by vascular remodeling. It is the final stage of chronic liver disease, mainly related to alcohol consumption, chronic viral infection, autoimmune, and metabolic etiologies [2].

Gastric cancer is presently the fourth most common cancer and the second most frequent cause of death in the world. Its symptoms are typically diagnosed when the disease is very advanced and in a metastatic phase [3]. Treatment for this disease is expensive and widely inaccessible in most parts of the world [4]. Consequently, liver cirrhosis (LCi) and gastric cancer (GCa) persist as a primary source of death despite decades of effort and expenditure on treatment [5].

The new concept in the management of patients with LCi and GCa should be prevention and early intervention to stabilize disease progression. Thus, the challenge is to develop effective and predictive methods.

Although much effort has been devoted to the development of effective therapeutic methods, there is still a need to improve premature detection methods. Early detection and accurate diagnosis of the onset of disease are the most promising approaches to accelerate the healing process. They could significantly reduce associated mortality.

Exhaled breath testing, as a non-invasive method and means of monitoring disease progression, has many advantages over other traditional methods. It guarantees easy sampling, real-time analysis, and potentially low cost [6]. In addition, several studies have even pointed to the volatile organic compounds (VOCs) emitted from a person's breath as indicators or markers of liver cirrhosis and stomach cancer [7,8]. For the purpose of analyzing these VOCs, various methods were employed, such as Gas Chromatography coupled with Mass Spectrometry (GC-MS) and Proton-Transfer-Reaction Mass Spectrometry (PTR-MS) [9,10]. Although these analytical methods allow the quantification and detection of various species at low concentrations, they have many disadvantages namely the complexity, expensiveness, and require skilled personnel. For these considerations, the development of advanced, low-cost gas sensors that are highly sensitive towards low VOC concentrations is of paramount concern. In this respect, electronic noses constitute an innovative method of VOC sampling, as these devices permit the online recognition of complex VOC mixtures by composite sensor arrays in conjunction with pattern recognition methods. [11]. They mainly follow an empirical approach, distinguishing "breathing patterns" arising from various VOCs by pattern recognition, which enables mixtures of gases to be distinguished independently of their individual molecular components [11].

In this research paper, the ability of an electronic nose (e-nose) system based on SnO_2 sensors to discriminate between three groups of patients with LCi and GCa versus HC was performed using exhaled breath. The e-nose data set was processed by employing principal component analysis (PCA), discriminant function analysis (DFA), and support vector machines (SVM).

2. Materials and Methods

2.1. Collection of Breath Samples

A total of 36 volunteers aged 26-77 years (both genders: 15 female & 21 male) were participated in this research work. These volunteers included 15 LCi, 8 GCa patients and 13 HC. Breath samples were collected according to the biomedical research ethics protocol of Avicenna University Hospital (Mohammed V University of Rabat) and after signed consent of each subject. The participants were asked to breathe into the Tedlar® bag in connection with mouthpiece before any beverages or food in-take and after mouth rinsing with purified water. Exhaled breath samples were collected in triplicate, and immediately transferred towards sensor arrays by pumping the content of each collected bag, for 10 minutes with flow rate of 200 mL/min.

2.2. E-nose measurements

The e-nose system used in these measurements comprised three main parts: unit of sampling composed by a Tedlar® bag related to a micro pump, a sensor array, and data acquisition unit. The sensor array comprises of five commercial gas sensors based on SnO_2: MQ-2, MQ-3, MQ-135, MQ-137 and MQ-138 (Hanwei Electronics Co. Ltd., Zhengzhou, China). A relative humidity sensor (Honeywell HIH 4000-002) and a temperature sensor (LM35DZ) from National Semiconductor were also included inside the sensor chamber to monitor the conditions of the experiment. Sensors responses were recorded employing a NI USB-6212 data acquisition board from National Instruments (Texas, USA). The e-nose system used in this study was described in detail in our previous work [12].

2.3. Data Pre-processing and Multivariate Analysis

Two sensing features were extracted from raw data of each sensor response:

Gs: Steady-state conductance calculated as the average value of the conductance change during the 9^{th} minute of breath measurement.

AUC: Area under the conductance curve in a time interval defined between 3^{th} and 9^{th} minute of breath measurement.

The dataset obtained from the e-nose was analyzed by pattern recognition methods.

PCA is a powerful, linear, and transform of multidimensional space that has been shown to be effective for the classification by e-nose system [13]. The aim of using this method is to allow a visual approach of the problem in a reduced representative space defined by principal components (PCs). The algorithm determines the first most important variance between samples, which is explained in the first principal component (PC1). The second largest variance is assigned to PC2 (the orthogonal to PC1). This will continue until all variance is explained in the multivariate data set. The percentage of the data variance contained in each principal component is given by the corresponding eigenvalue.

DFA is one of the most frequently used supervised pattern recognition method. The purpose of using the DFA method is to predict the classification of a sample into predefined groups [14]. DFA procedure is based on the determination of discriminant functions, which minimize the ratio of inside-groups, called intra-class variance and maximize the ratio between groups, called inter-class variance.

SVMs is a non-linear supervised recognition method based on the notion of maximum margin, which is the distance between the separation boundary and the nearest samples called support vectors [15]. It is obvious that there is a multitude of valid hyperplanes but the remarkable property of SVMs is that this hyperplane must be optimal. We therefore look for the one among the valid hyperplanes that passes "in the middle" of the points of the two classes. The optimal separation hyperplane is the one that maximizes this margin. The SVM technique was originally developed for two-class classification. However, in most practical problems, there is no linear separation between the data. To overcome this limitation, two types of SVM approaches can be considered. The first, which is used in this paper, consists of constructing and combining several one-against-one or one-against-all binary classifiers. In the second approach, all data are directly considered in a single optimization formulation.

3. Results and discussion

3.1. E-nose responses

Figure 1 displays the MQ-138 gas sensor responses towards exhaled breath samples from patients with LCi, GCa, and HC. As remarked, the sensor response of LCi patient is relatively higher than the GCa patient and HC. Furthermore, the sensor response corresponding to breath sample of GCa was evidently higher to HC. This behaviour can be justified by differences of breath VOCs concentrations of HC compared to LCi and GCa patients [16,17].

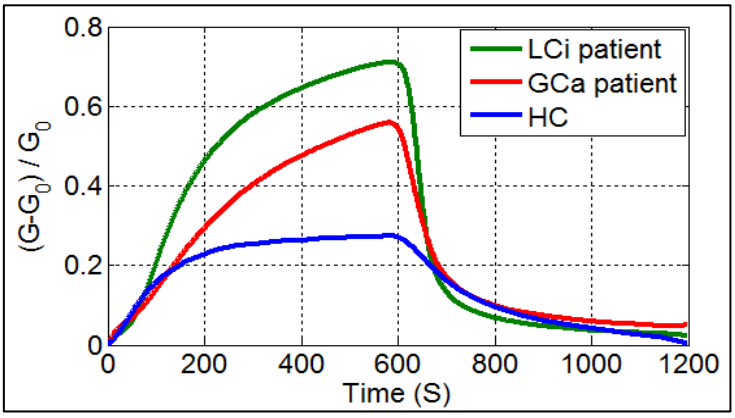

Figure 1. Sensor responses in presence of three breath samples from LCi, GCa patients, and HC using MQ-138 sensor.

The radar plots with unitary radius are described in the Figure 2, which shows an illustrative case in order to see if there are differences and/or similarity in patterns (i.e.

breath-prints) between breath samples corresponding to LCi, GCa patients and HC. Indeed, a clear pattern variation between the breath-prints of different studied cases is noticed. The size of breath-print of HC was found smaller compared to breath-prints of LCi and GCa patients. This behavior can be justified by an increase of some breath VOCs concentrations in LCi and GCa patients than in HC breath [16,17].

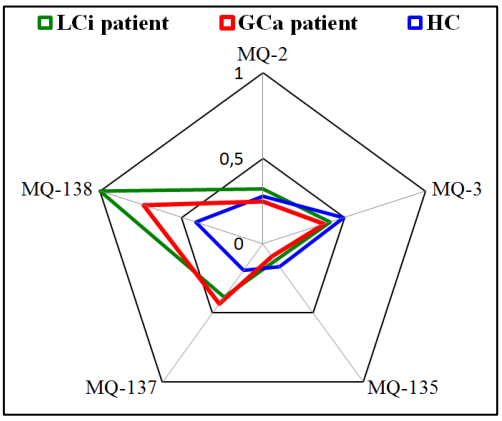

Figure 2. Radar plots of the five gas sensors of e-nose towards three-exhaled breath from patients with LCi, GCa, and HC expressed by Gs as feature.

3.2. PCA classification results

PCA plot was employed as an exploratory method to investigate the capability of e-nose device to discriminate between LCi, GCa patients and HC. The PCA has shown patterns with a good discrimination of the studied breath samples corresponding to LCi, GCa patients and HC with total variance of 96.42%, as can be seen in Figure 3. This behaviour can be explained by the different types of VOCs that exist in the breath samples corresponding to LCi, GCa patients, and HC [16,17]. Therefore, data processing using the PCA technique has demonstrated the effectiveness of the offered e-nose to differentiate between three studied groups basing on their health status.

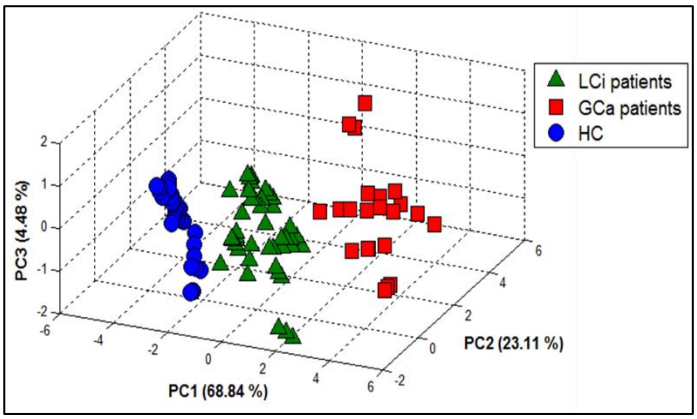

Figure 3. Unsupervised PCA plot displaying data-points of breath samples related to the three-health states with data gathered from e-nose.

3.3. DFA classification results

DFA was also performed to assess the e-nose feasibility to cluster different exhaled breath patterns corresponding to LCi, GCa patients and HC. Figure 4 shows the first two DFA functions for classification of exhaled breath samples using e-nose dataset. As a result, it is found that the breath samples are well clustered according to their health states. These findings are in good agreement with PCA results, which support our training model.

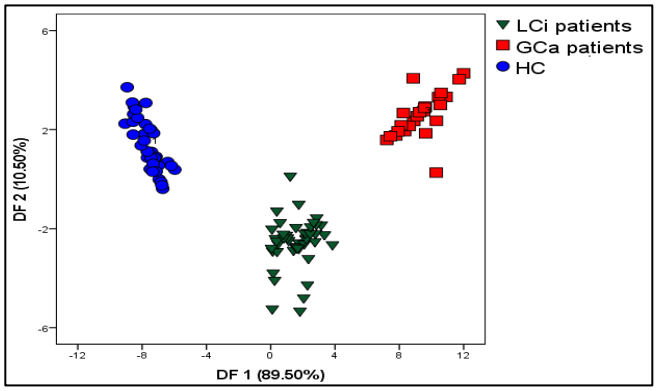

Figure 4. Supervised DFA plot exhibiting data-points of breath samples related to the three-health states with data gathered from e-nose.

3.4. SVMs classification results

SVMs method was employed as a supervised recognition method to test the ability of the e-nose in matter of discrimination among LCi, GCa patients and HC.

Table 1 represents the SVMs classification results of 108 exhaled breath measurements of LCi, GCa patients and HC. Rows indicate actual categories and columns the predicted ones. The SVMs method reached a 100% success rate for the recognition of the three studied cases. In the light of this result, it can be concluded that e-nose system is capable to distinguish between breath VOCs of LCi, GCa patients and HC.

Table 1. SVMs results for the classification of 108 breath samples regarding their health states by using the e-nose system with a success rate of 100%.

Actual	**Predicted**		
	LCi patients	**GCa patients**	**HC**
LCi patients	45		
GCa patients		24	
HC			39

4. Conclusion

This work has revealed the potential of e-nose system to distinguish between LCi, GCa patients and HC. The e-nose responses displayed a clear variation and significant differences between three studied health states. Furthermore, PCA and DFA methods have allowed differentiating between LCi, GCa patients and HC. In addition, a classification success rate of 100% was achieved by SVMs using one-against-one approach for LCi, GCa patients and HC. In the light of these results, the e-nose device based on five tin dioxide (SnO_2) sensors combined with pattern recognition methods has proven to be a useful tool for differentiating three studied groups. Therefore, this technology has the capability of non-invasively and painlessly detecting diseases. It could be extended in clinical practice for diseases diagnosis.

Acknowledgments: Authors gratefully acknowledge Moulay Ismaïl University of Meknes for financial support of the project "Research support".

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, Q.; Lee, W.Y. Urinary metabolites for urological cancer detection: a review on the application of volatile organic compounds for cancers. *Am. J. Clin. Exp. Urol*. **2019**, 7, 232.
2. Pijls, K.E.; Smolinska, A.; Jonkers, D.M.; Dallinga, J. W.; Masclee, A.A.; Koek, G.H.; van Schooten, F.J. A profile of volatile organic compounds in exhaled air as a potential non-invasive biomarker for liver cirrhosis. *Sci. Rep*. **2016**, 6(1), 1-8.
3. Daryabari, S.S.; Safaralizadeh, R.; Hosseinpourfeizi, M.; Moaddab, Y.; Shokouhi, B.; Gastrointest, J. Overexpression of SSH1 in gastric adenocarcinoma and its correlation with clinicopathological features. *J. Gastrointest. Oncol*. **2018**, 9, 728–733.
4. Mohammadi, M.; Zarghami, N.; Hedayati, M.; Ghaemmaghami, S.; Yamchi, R.M. ; Mohaddes M. Visfatin effects on telomerase gene expression in AGS gastric cancer cell line. *Indian J. Cancer* **2015**, 52, 32–35.
5. Kumar, S.; Huang, J.; Abbassi-Ghadi, N.; Mackenzie, H.A.; Veselkov K.A.; Hoare, J.M.; Hanna, G.B. Mass spectrometric analysis of exhaled breath for the identification of volatile organic compound biomarkers in esophageal and gastric adenocarcinoma. *Ann. Surg*. **2015**, 262, 981–990.
6. Lourenço, C.; Turner, C. Breath analysis in disease diagnosis: methodological considerations and applications. *Metabolites*. **2014**, 4(2), 465-498.
7. Durán-Acevedo, C.M.; Cáceres-Tarazona, J.M. Low-cost desorption unit coupled with a gold nanoparticles gas sensors array for the analysis of volatile organic compounds emitted from the exhaled breath (gastric cancer and control samples). *Microelectron. Eng*. **2020**, 237, 111483.
8. Rondanelli, M.; Perdoni, F.; Infantino, V.; Faliva, M. A. ; Peroni, G.; Iannello, G.; Nichetti, M.; Alalwan, T. A.; Perna, S.; Cocuzza, C.Volatile organic compounds as biomarkers of gastrointestinal diseases and nutritional status, *J. Anal. Methods Chem*. **2019**, 2019, 1-15.
9. Buszewski, B.; Ligor, T.; Jezierski, T.; Wenda-Piesik, A.; Walczak, M.; Rudnicka, J. Identification of volatile lung cancer markers by gas chromatography–mass spectrometry: comparison with discrimination by canines. *Anal. Bioanal. Chem.* **2012**, 404, 141–146.
10. Wehinger, A.; Schmid, A.; Mechtcheriakov, S.; Ledochowski, M.; Grabmer, C.; Gastl, G.A.; Amann, A. Lung cancer detection by proton transfer reaction mass-spectrometric analysis of human breath gas. *Int. J. Mass Spectrom*. **2007**, 265, 49-59.
11. Dragonieri, S.; Schot, R ; Mertens, B. J.; Le Cessie, S.; Gauw, S. A.; Spanevello, A.; Sterk, P. J. An electronic nose in the discrimination of patients with asthma and controls. *J. Allergy Clin. Immunol*, **2007**, 120(4), 856-862.
12. Saidi, T.; Zaim, O.; Moufid, M.; El Bari, N.; Ionescu, R.; Bouchikhi, B. Exhaled breath analysis using electronic nose and gas chromatography–mass spectrometry for non-invasive diagnosis of chronic kidney disease, diabetes mellitus and healthy subjects. *Sens. Actuators B Chem.* **2018**, 257, 178-188.
13. Aït-Sahalia, Y.; Xiu, D. Principal component analysis of high-frequency data. *J Am Stat Assoc*. **2019**, 114(525), 287-303.
14. Bougrini, M.; Tahri, K.; Haddi, Z.; Saidi, T.; El Bari, N.; Bouchikhi, B. Detection of adulteration in argan oil by using an electronic nose and a voltammetric electronic tongue. *J. Sens*. **2014**, 2014, 1-10.
15. Guenther, N.; Schonlau, M. Support vector machines. *Stata J.* **2016**, 16(4), 917-937.
16. Ghosh, C.; Singh, V.; Grandy, J. Recent advances in breath analysis to track human health by new enrichment technologies. *J. Sep. Sci*. **2020**, 43, 226–240.
17. Rondanelli, M.; Perdoni, F.; Infantino, V. Volatile organic compounds as biomarkers of gastrointestinal diseases and nutritional status. *J. Anal. Methods Chem.* **2019**, 2019, 1-15.