IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 18 February 2025, accepted 5 March 2025, date of publication 11 March 2025, date of current version 19 March 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3550453

RESEARCH ARTICLE

Lung Cancer Detection Utilizing Mixed Sensor Based Electronic Nose

UMIT OZSANDIKCIOGLU, AYTEN ATASOY, AND YUSUF SEVIM
Department of Electrical and Electronics, Faculty of Engineering, Karadeniz Technical University, 61080 Trabzon, Türkiye

Corresponding author: Umit Ozsandikcioglu (umitozsandikci@ktu.edu.tr)

This work was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) and TÜBİTAK Marmara Research Center Materials Sciences Institute under Project 215E380.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Karadeniz Technical University under Application No. 24237859.

ABSTRACT The lungs are the most important organ of the respiratory system. The volatile organic compounds found in breath, which usually originate from the blood and enable us to observe different processes in the body, are carried to the lungs through the blood and then exhaled by breath. In this study, a mixed sensor based electronic nose circuit was developed using eight metal oxide semiconductors and 14 Quartz Crystal Microbalance gas sensors. A total of 100 volunteers participated in the study, including 20 healthy volunteers who did not smoke, 20 healthy volunteers who smoked, and 60 lung cancer volunteers. Throughout this study, 338 experiments were conducted using breath samples. Data dimension reduction was achieved using linear discriminant analysis and principal component analysis algorithms. The individual classification accuracies for the metal oxide semiconductor and quartz crystal microbalance sensor data are 81.54% and 73.18%, respectively. Upon combining the sensor data, a noticeable increase in accuracy of 85.26% was observed. In this study, the performance of the developed system was enhanced using principal component and linear discriminant analyses. While the highest classification accuracy increased to 88.56% with the feature matrix obtained using the principal component analysis method, this value was obtained with 94.58% accuracy using the linear discriminant analysis method.

INDEX TERMS Breath analysis with type different sensor based electronic nose, lung cancer diagnosis, data processing, dimension reduction and classification.

I. INTRODUCTION

A. MOTIVATION

Lung Cancer (LC) is the deadliest type of cancer in the world. According to data from the World Health Organization (WHO), the global mortality rate attributed to LC exceeds the combined mortality rates of kidney, colorectal, and prostate cancers [1], [2]. Despite ranking second after colorectal cancer in terms of diagnosis rates, LC remains the deadliest form of cancer. Globocan 2020 reports indicate that there are 19.3 million new cancer cases and 10 million cancer-related deaths worldwide. Figure 1 presents the global distribution of new cancer cases and cancer-related deaths in 2020, as reported by Globocan [3]. Research indicates that the 5-year survival rate for individuals diagnosed with early stage LC can increase to as high as 90%. However, this rate declines significantly to below 10% for those whose cancer is identified at an advanced stages [4]. The variations in the 5-year survival rates based on the LC stage at diagnosis are illustrated in Figure 2 [5]. Because LC typically presents with few symptoms in the early stages, medical imaging methods are the most commonly used tools for detecting tumors. Survival rates are significantly higher for individuals whose cancer is detected early enough to permit surgical intervention [6]. For the early diagnosis of LC, people in the high-risk category were analyzed using various medical techniques. These techniques include common combinations of saliva cytology [7], [8], tumor biomarkers circulating within the human body [9], [10], protein structure of the blood [11], [12], chest tomography [13], [14],

The associate editor coordinating the review of this manuscript and approving it for publication was Vincenzo Conti.

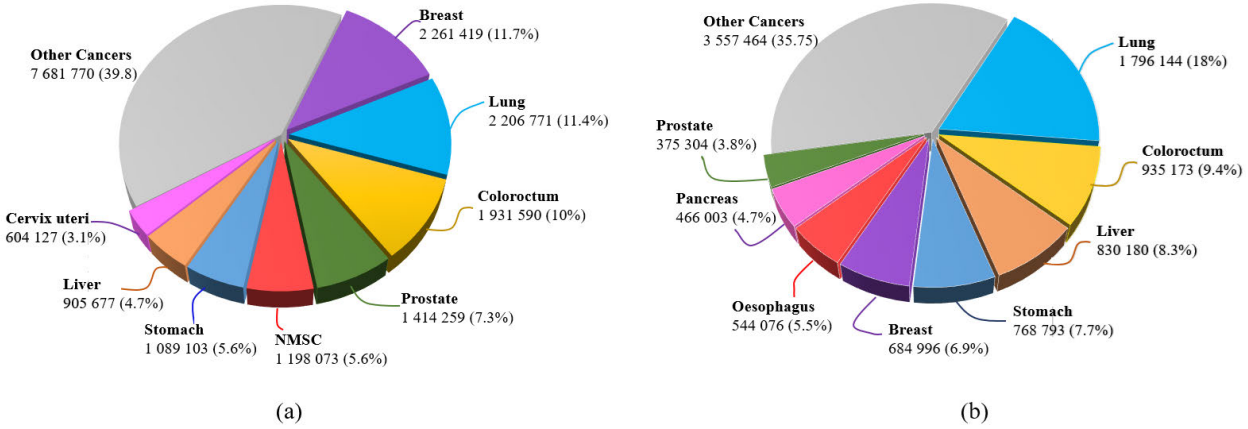

FIGURE 1. The number of new cancer cases (a), cancer related deaths in 2020 (b).

nuclear magnetic resonance [15], [16], chest radiography, and magnetic resonance imaging chest radiography and magnetic resonance imaging, such as low-dose computed tomography [17], [18]. Given the limited capabilities of these methods, they have been utilized for several years. In terms of diagnosis, studies have improved their performance. Chest radiography (CXR), a medical imaging method, has not been able to achieve a notable reduction in the mortality rate of patients with LC over the past few decades [19], [20]. Although this technique has low sensitivity, it also results in a significant occurrence of false negative results. Furthermore, the elevated radiation intensity associated with CXR restricts their application in this domain. In a study carried out by Mazzone et al., They indicated that computer-assisted tomography (CT) can address certain limitations of the conventional CXR method; however, there remain areas that require further enhancement [21]. CT generates a substantial number of false negative results and fails to detect numerous cases of early stage LC. Furthermore, owing to its ability to detect lesions smaller than 1 mm in diameter, CT yields a substantial number of false-positive outcomes, leading to unnecessary biopsies [22]. To mitigate this issue, verification was performed using positron emission tomography (PET) scanning [23]. However, the high costs associated with CT and PET restrict their utilization for early detection of LC. The fact that the imaging methods used are not very successful in detecting early stage LC and the reduction in mortality rates has led scientists to conduct studies in the field of breath analysis.

B. PROPOSED APPROACH

Breath analysis aims to identify and measure volatile organic compounds (VOCs), which may serve as disease markers, in the breath. Chemical compounds formed as a result of organ function are transferred from organs to the blood [24], [25]. Blood is cleaned in the lungs, and some chemical compounds in the blood are transferred out of the body through breathing [26]. While some of the chemical compounds found in human breath originate from the environment in which individuals live, some emerge as a consequence of the overall functioning of the human organs [27]. In a 1970 study, Pauling et al. analyzed the composition of human breath and identified more than 200 chemical compounds using gas chromatography [28]. Following this groundbreaking study by Pauling et al., scientists have continued their research in the field of breath analysis. Over the past decade, the composition of human breath has been broadly characterized, particularly with the aid of advanced medical electronic devices [29]. Based on these studies, it was concluded that human breath primarily contains three types of chemicals. These chemicals are inorganic compounds, VOCs, and non-volatile compounds [23]. Recent studies have estimated that 3000 VOCs are present in the gas phase of human breath [30]. These VOCs are produced in the human body, including, unsaturated and saturated hydrocarbons [31], [32], and compounds that include oxygen [33], sulfur, and nitrogen [34], [35]. VOCs transferred to the human breath through the blood provide important information, especially regarding the biochemical functioning of the human body. When breath analysis studies were examined, it was determined that in cases of febrile illness and sepsis, the pentane concentration in the breath was higher than that in healthy individuals [36], [37]. Acetone and methyl nitrate are present in elevated concentrations in the breath of individuals with diabetes, a prevalent global condition [38], [39].

Sulfur-containing compounds (Dimethyl Sulfide, ethyl, and methyl mercaptans) are generally formed because of abnormalities within the body's excretory system [40], [41]. Consequently, people with renal disease have higher levels of sulfur-containing chemicals in their breaths. Di/trimethylamine and ammonia are nitrogen-containing VOCs that are more prevalent in the breath of individuals with kidney illness [42]. In a study examining LC, which is the subject of this study, Peng et al. examined the VOCs found in the breath of cancer patients using gas chromatography-mass spectrometry (GC-MS) in four different types of cancer: LC, bowel cancer, breast cancer, and prostate cancer. They identified 80% non-overlapping VOCs in the breath of

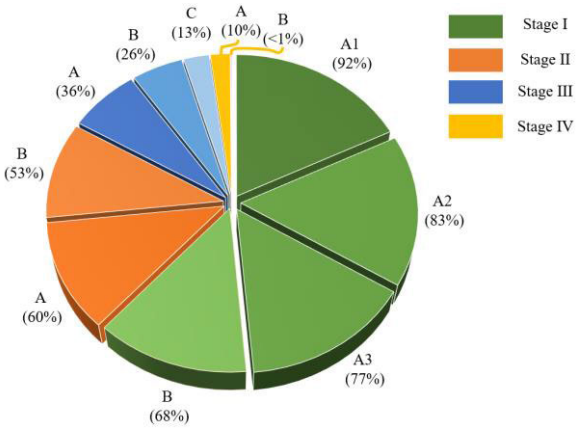

FIGURE 2. Non-small cell LC 5-year survival rates according to the stage of cancer.

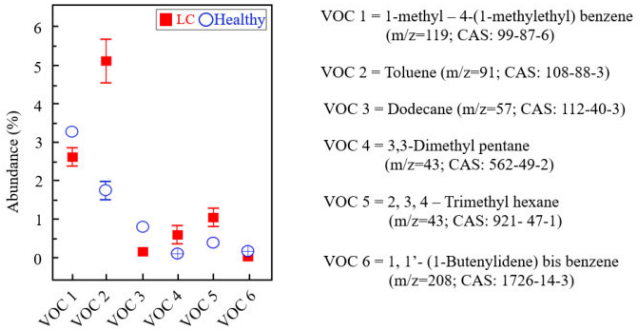

FIGURE 3. The most suitable VOCs were selected for distinguishing between healthy subjects and patients suffering from LC.

both healthy individuals and patients. There were 34 VOCs identified for bowel cancer, 33 for LC, 36 for prostate cancer, and 54 for breast cancer. The distribution of the most appropriate VOCs determined in their study to distinguish between healthy volunteers and LC patients is shown in Figure 3 [43]. Poli et al. investigated aldehyde levels in the exhaled breath of patients with LC. The quantities of aldehydes, specifically propanal, butanal, pentanal, hexanal, heptanal, octanal, and nonanal, varied significantly among the breath samples from the different collection groups [44]. Jia et al. examined 25 studies in the literature in their review article on the breath content of patients with LC. As a result of their examination, four or more studies identified VOCs in the breath of LC patients at significantly different levels compared to healthy individuals. These VOCs were identified as Propanol, Pentanol, Isoprene, Hexane, Acetone, Toluene, Ethanol, Benzene, Butanol, Ethyl Benzene, Pentanal, Heptane, Styrene, 2- Butanol, Propanal [45]. The GC-MS approach is mostly utilized in these studies to ascertain breath content. In addition to these methods, ion flow tube mass spectrometry, laser absorption spectrometry, and infrared spectroscopy are frequently employed to identify VOCs in breath. However, the methods required for these methods are quite expensive, and the tests are performed for a long time. Additionally, these devices require the pre-concentration of breath samples to enhance the success of VOC detection [46], [47].

C. HARDWARE APPROACHES

Electronic nose (e-nose) is a device inspired by the biological olfactory system. Owing to its structure, having both a chemical detection system and a data processing system, it can detect many simple or complex odors. These devices comprise three major components: an electronic unit, pattern recognition unit, and sensor unit [30], [48]. A comparison of the e-nose and mammalian olfactory systems is shown in Figure 4. Chemical sensors are the basic components of a sensor unit, which is one of the most basic components of e-nose circuits. These sensors generate electrical signals depending on the properties of the chemicals applied to them. The sensor types mostly used in e-noses over the last 50 years can be used as optical gas sensors, polymers, metal oxide semiconductors (MOS), and field effect transistors (FET), catalytic, electrochemical, and piezoelectric sensors [50]. MOS sensors, one of the sensors used in this study, are among the most frequently used in e-nose applications. Materials such as SnO2, ZnO, CuO, TiO2, WO3, and NiO are generally used as sensing chemicals in these sensors. In MOS sensors, the sensor material reacts with the target gas at a suitable temperature, and, as a consequence of this chemical reaction, changes the conductivity of the sensor [51], [52].The alteration in the sensor conductivity is transformed into electrical signals using suitable circuits and subsequently transmitted to the computer. Literature indicates that MOS sensors are utilized in both commercially available and custom-built electronic noses [53], [54], [55].

Quartz crystal microbalance (QCM) sensors are another popular gas sensor for e-noses. These sensors were composed of a quartz disk sandwiched between two parallel gold electrodes. Quartz, a piezoelectric material, undergoes mechanical deformation upon application of an alternating current to the electrodes. Additionally, a standing wave is produced in the resonator if the current frequency matches the resonant frequency of the QCM sensor. The mass of the sensor has an impact on the standing wave. The target gases were detected using a thin layer covering the QCM sensors. The thin sheet absorbed the particles of the target gas by, changing the resonant frequency of the sensor. An appropriate electronic circuit calculates and sends this frequency information to a computer. Useful data regarding the target gases can be obtained using this frequency deviation [56], [57], [58].

D. DATA ANALYSIS APPROACHES

In the literature, there are important studies carried out with an e-nose on the detection of various illnesses using breath analysis [59], [60]. Gasparri et al. utilized an e-nose equipped with QCM sensors in their research.146 volunteers, 76 healthy and 70 LC patients, participated in the study. In their study, the sensor data were used as raw data without preprocessing. Partial Least Squares Discriminant (PLSD) analysis was used as a classifier, and a sensitivity of 81% was

achieved [61]. In another study involving 43 LC patients at different stages and 58 healthy volunteers, Blatt et al. utilized an e-nose equipped with MOS sensors. The researchers used Principal Component Analysis (PCA) and non-parametric Linear Discriminant Analysis (LDA) for data size reduction. k-nearest neighbor (k-NN), Artificial Neural Networks (ANN), and linear and quadratic discriminant function-based classifiers were used as the classification algorithms. In this study, researchers obtained an accuracy of 92.6% [62]. Nakleh et al. conducted experiments using an e-nose with breath samples collected from patients with LC (n=45), other cancer types (n=357), and control subjects (n=411). They employed an e-nose based on gold (Au) and single-walled carbon nanotubes, achieving a classification success rate of 85% [63]. Binson et al. attempted to differentiate LC patients from others by using an e-nose equipped with MOS sensors. The researchers achieved a sensitivity, specificity, and accuracy of 91.3%, 84.4%, and 94.4%, respectively [64]. Mazzone et al. used a system containing a colorimetric sensor array in their study, in which 229 volunteers, 92 LC patients, and 137 healthy individuals participated. In this study, which aimed to distinguish between LC and healthy volunteers, an accuracy of 81.1% was obtained [21]. In their study involving 145 volunteers, Van der Goor et al. attempted to distinguish data from individuals with LC from those with benign tumors. The data obtained from the experiments carried out with the Aeonose brand ready-made e-nose device were classified using an ANN algorithm. As a result of this classification, 83% sensitivity and 84% specificity values were achieved [65]. Saidi et al. analyzed the breath samples of volunteers in the LC (n=32) and control (n=12) groups using a chemical gas sensor-based e-nose. The researchers achieved a classification accuracy of 98.6% in their study [66]. The literature reveals that nearly all e-noses utilized in LC detection studies rely on a single type of gas sensor, including both commercially available e-noses and those specifically designed by researchers [67]. In this study, a hybrid sensor-based e-nose system was developed that incorporated eight MOS sensors and 14 QCM sensors. In this manner, breath samples of the volunteers were converted into both conductivity and frequency information. Utilizing data from both sensors significantly enhanced the system performance, as demonstrated by the results. Thus, an inexpensive electronic circuit that can detect LC with high accuracy using breath analysis was developed in this study.

II. MATERIAL AND METHODS

This section begins with a comprehensive explanation of the e-nose system that was developed in this study. Following this, the processes of data preprocessing, feature extraction, and pattern recognition were elaborated.

A. EXPERIMENTAL CONFIGURATION

A schematic representation of the e-nose circuit, in which breath analyses were performed throughout this study, is shown in Figure 5. The experimental setup consisted

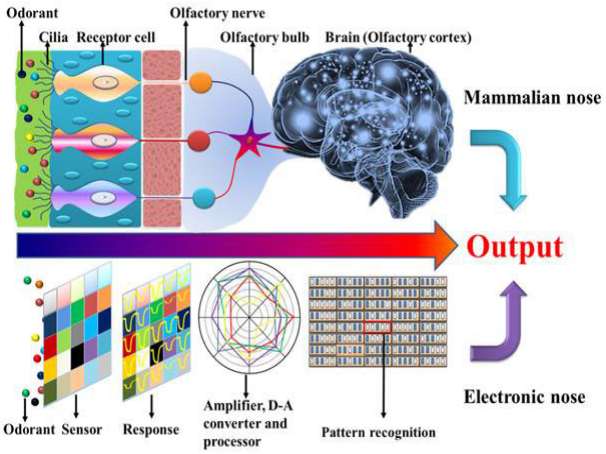

FIGURE 4. An illustrative diagram displaying the structure of the mammalian nose in comparison with e-nose [49].

of five components. These components include, dry air tubes, Teflon pipes, vacuum pumps, solenoid valves, QCM sensor chambers, MOS sensor chambers, sensor interface circuits, relay cards, analog-digital data acquisition cards, and computers. The MOS sensors used in this study were purchased from among the sensors produced by Figaro. For the selection of MOS sensors, VOCs present at different concentrations in the breath of healthy individuals and LC patients were searched for in the literature. After this search, we decided to use the TGS813, TGS816, TGS826, TGS832, TGS2602, TGS2610, TGS2612, and TGS2620 gas sensors. The MOS sensors in the sensor chamber and target gases of the sensors are shown in Figure 6. The QCM sensor is the second type of sensor employed in e-nose circuits. The QCM sensors used in this study were produced by the TUBITAK Marmara Research Center, located in Gebze, Kocaeli, Turkey. The team at TUBITAK Marmara Research Center used AT-cut quartz crystals supplied by Klove Electronics (Westerlo, Belgium). The QCM sensors were cleaned of any remaining solvent using a dry air stream during the production phase, and the chambers housing the QCM sensors were equipped with an internal frequency counter circuit. This circuit detects the frequency values generated by the QCM sensors and transmits them to the computer. Thus, the frequency data corresponding to breath samples were obtained. Figure 7 and Figure 8 present the images of the QCM sensor chambers and their block diagrams, respectively.

Throughout this work, the entire system was cleaned with dry air composed of 21% oxygen and 79% nitrogen, both before and after each experiment. The flow rate of the dry air was maintained at10 L/min. This cleaning process, which included sensors and Teflon pipes to facilitate gas flow, ensured stable and reliable sensor data. Teflon pipes were utilized to transport both dry air and breath samples, as they are chemically inert and do not retain odor molecules, thus preventing contamination from previous experiments. A 12Volt direct current vacuum pump was employed to

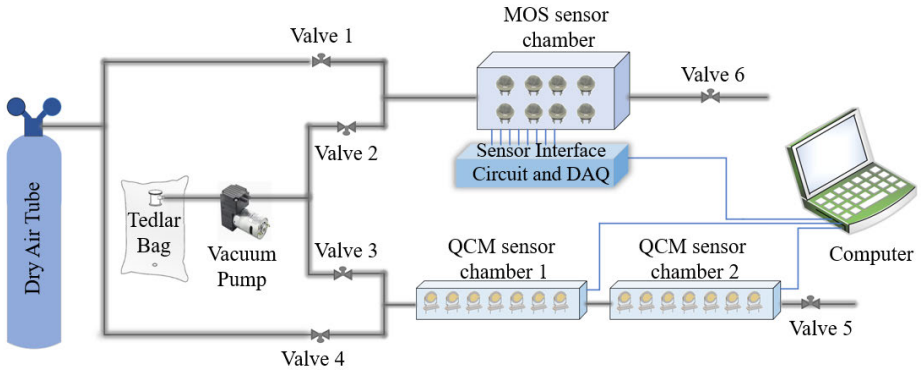

FIGURE 5. Block diagram of the developed e-nose system.

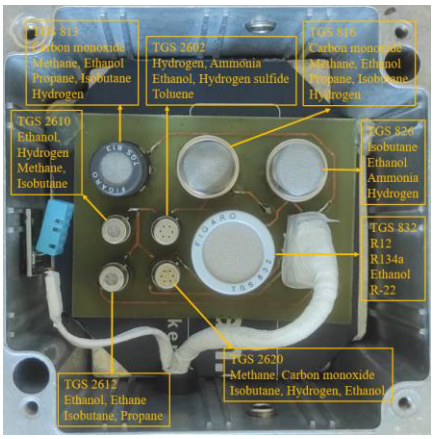

FIGURE 6. MOS sensors and their target gases.

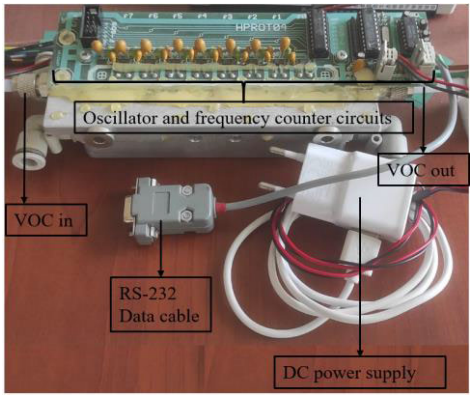

FIGURE 7. QCM sensors chamber.

deliver breath samples to the sensors. Use of a constant supply voltage for the vacuum pump ensured that the collected breath was introduced to the sensor at a uniform flow rate during all the experiments. During the system setup phase, the supply voltage was determined based on the capacity of the Tedlar bags and sensor chambers, resulting in a flow rate of 3.6 liters/min for breath sample delivery. Six solenoid valves were incorporated into the experimental setup to direct breath samples and dry air appropriately. These valves, manufactured by JELPC (Ningbo, Zhejiang, China), operate at 24V DC with a power consumption of 4.8 W. During the experimental process, the transfer of breath samples and dry air within predetermined durations and directions was facilitated by vacuum pumps and solenoid valves in the experimental setup. To transmit breath samples and dry air to the relevant areas in the developed e-nose, the valves on the transmission line must be in an open position. To allow gas flow, the valves used in this study require a supply of 24 V DC voltage. Similarly, for the breath samples to be drawn from the Tedlar bags and delivered to the sensor chambers, a 12 V DC voltage was supplied to the vacuum pump. The required voltages of 12 V and 24 V in this study were supplied to the valves and vacuum pump through the driver circuit, as illustrated in Figure 9. In this driver circuit, a 5 V DC voltage obtained from the digital outputs of the USB-6218 multifunctional I/O device (National Instruments) is applied to the base terminal of transistor T1. Consequently, the transistor switches to the conduction state, allowing the current to pass through the relay coil connected to its collector terminal, causing the relay to switch from the normally closed (NC) to the normally open (NO) position. Consequently, the 24 V voltage present at the relay's NO contact was transferred to the valves, while a 12 V voltage was supplied to the vacuum pump. To execute the experimental procedures, necessary calculations were performed, and a control board for six valves and one vacuum pump was designed, as shown in Figure 10. The electronic schematic of this circuit was designed using Proteus v8.6 software and was physically realized by our team. The inputs of the developed control board were provided via the USB-6218 multifunctional I/O device (National Instruments) and programmed using MATLAB 2022b. This programming ensured that all the experimental steps were executed automatically in a pre-defined sequence and duration. The experiments conducted using the established e-nose system comprised of four distinct stages. During the preliminary stage, the entire e-nose was purged using dry air for 130 s, after which all valves were closed. In the second stage, the vacuum pump transported the collected breath to the sensor chambers for more than 40 s. In the third stage, the valves were closed for 30 s to allow the sensors to fully react with breath samples. Finally, the e-nose was purged with dry air for 140 s to ensure that it was

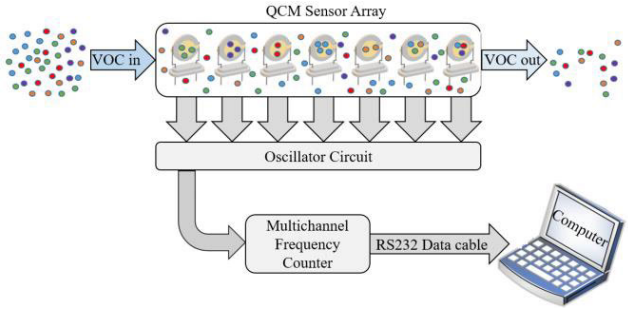

FIGURE 8. QCM sensor chamber block diagram.

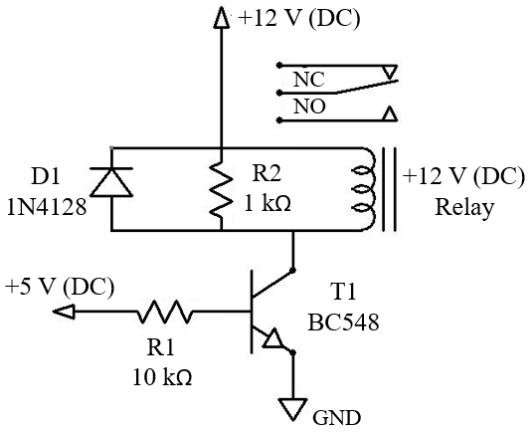

FIGURE 9. QCM sensors chamber.

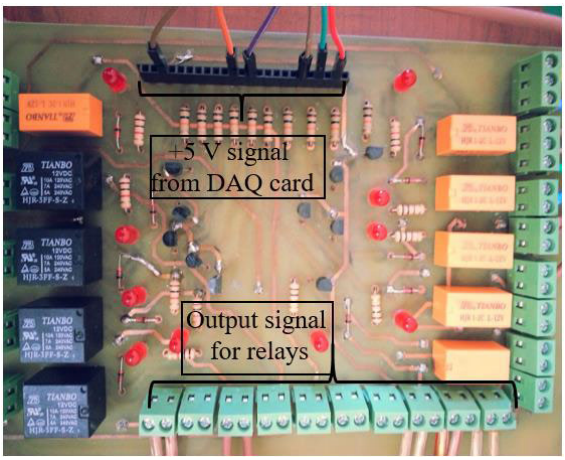

FIGURE 10. QCM sensors chamber.

prepared for subsequent experiments. Figure 11 provides a visual representation of the experimental setup developed in this study. A representative data sample from the MOS and QCM sensors is shown in Figure 12.

B. BREATH SAMPLES COLLECTION

This work involved 338 experiments, utilizing breath samples collected from 60 volunteers with LC who were treated at the Hospital of the Faculty of Medicine, Karadeniz Technical University, and 40 healthy volunteers. Before initiating the study, all necessary documents were obtained from the Ethics Committee of Karadeniz Technical University,

FIGURE 11. Developed e-nose circuit.

and the assigned ethics approval number was 24237859-517. After LC diagnosis in each volunteer, participants in the study were confirmed according to bronchoscopy and biopsy results, and breath samples were collected from the patients. The tumor stages of the lung cancer patients participating in the study were evaluated using the 7th edition of the TNM classification system, which assesses Tumor Extent (T), Regional Lymph Node Metastases (N), and Distant Metastasis (M), as defined by the American Joint Committee on Cancer (AJCC).

Prior to the collection of breath samples, patients did not undergo any surgical procedures. The study participants were healthy volunteers with no pregnancy status or any other lung disease. Healthy volunteers who smoked constituted half of all healthy volunteers. All volunteers participating in the study voluntarily exhaled. Information on all volunteers for whom breath samples were collected is presented in Table 1. The breath of the individuals participating in the experiments was collected using 5-liter Tedlar bags (CEL Scientific, Cerritos, CA, USA). Because Tedlar bags do not allow gas transfer between the internal and external environments, experiments can be performed without any change in the content of the breath samples. The Tedlar bags used to collect breath samples had a polypropylene connection piece that determined the direction of exhalation.

A visual diagram of the Tedlar bag and polypropylene connection piece is shown in Figure 13. Breath samples were collected by exhalation in the direction indicated by number 1. If the breath sample is to be filled in the bag in the direction indicated by number 2, and the pipeline indicated by number 3 must be closed. If this pipeline is open, the air exhaled in direction 1 is directed out of the bag using path number 3. A timing method was applied to ensure that alveolar air was collected instead of the upper respiratory tract. In this study, it was assumed that the duration of deep breathing was approximately 6 s. Accordingly, the pipeline indicated by number 3 was kept open for the first 3 s, whereas the volunteers exhaled and then closed. Thus, the first part of the breath, which may contain external factors or contamination, was expelled, and only the air coming from the lungs was

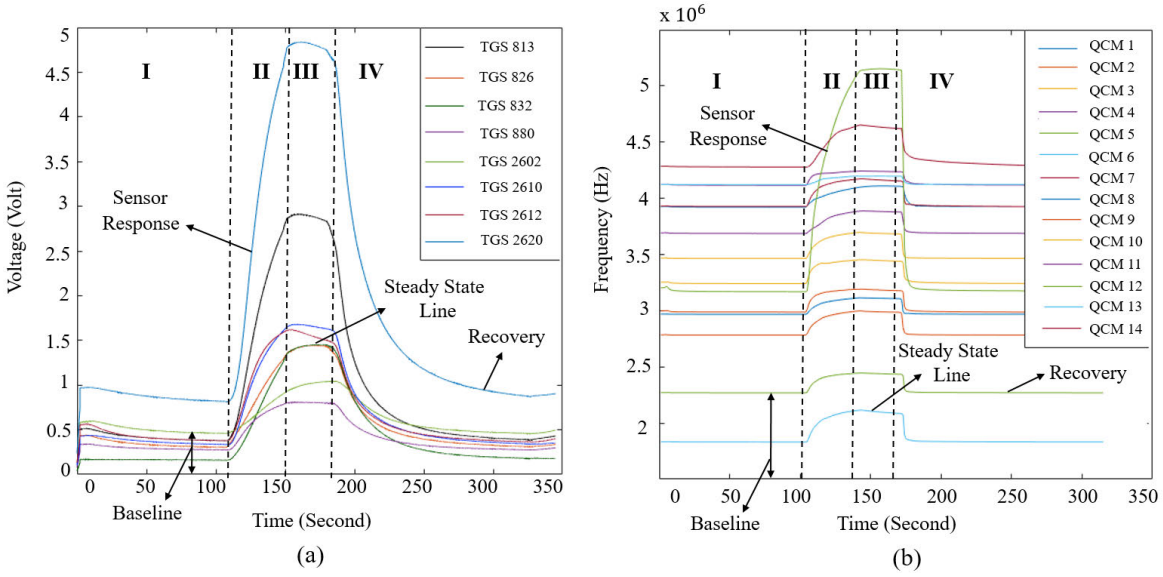

FIGURE 12. The signals of MOS sensor (left) and QCM sensor (right) generated as a result of an experiment.

TABLE 1. Informative data of the whole volunteers.

	Volunteer with LC (60 Person)	Healthy volunteer (40 Person)
Age (mean/std. deviation)	60.7 / 8	48.2 / 9
Gender (female/male)	13/47	10/30
Smokers/Nonsmokers	0/60	20/20
Exsmokers	45	0

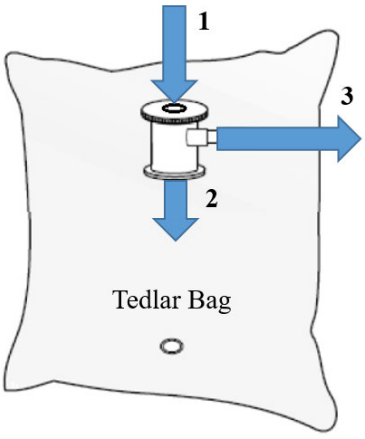

FIGURE 13. The visual diagram of tedlar bag and polypropylene pipeline.

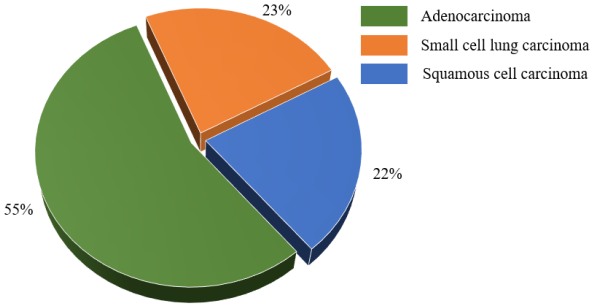

FIGURE 14. The allocation of patients by histological cancer type.

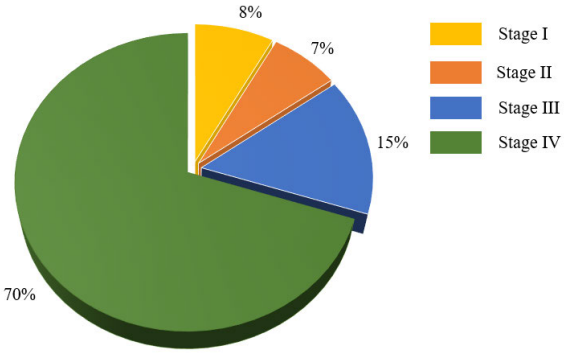

FIGURE 15. The allocation of patients by stage of cancer.

sampled. Volunteers were required to fast overnight and provide breath samples during fasting to participate in the experiments. Volunteers refrained from smoking for 2 h before the test. Figure 14 and Figure 15 show the cancer stages and cell types, respectively, of the volunteers included in this study.

C. FEATURE EXTRACTION AND DATA ANALYSIS

Before feature extraction, pre-processing was applied to the data. To accurately visualize the data and obtain useful information, reference correction was applied to the data obtained from both the MOS and QCM sensors using equations (1) and (2), respectively.

$$v_{n,s,r}(t) = v_{n,s}(t) - v_{n,s}(0) \quad (1)$$

$$f_{n,s,r}(t) = f_{n,s}(t) - f_{n,s}(0) \quad (2)$$

In equation (1), $v_{n,s}(t)$ is the raw sensor data, $v_{n,s}(0)$ is the sensor data obtained at the first moment when the collected breath is delivered to the sensors, and $v_{n,s,r}(t)$ is the reference corrected sensor data. A representative data sample from the reference-corrected MOS sensor data and raw MOS sensor data is shown in Figure 16. In equation (2),

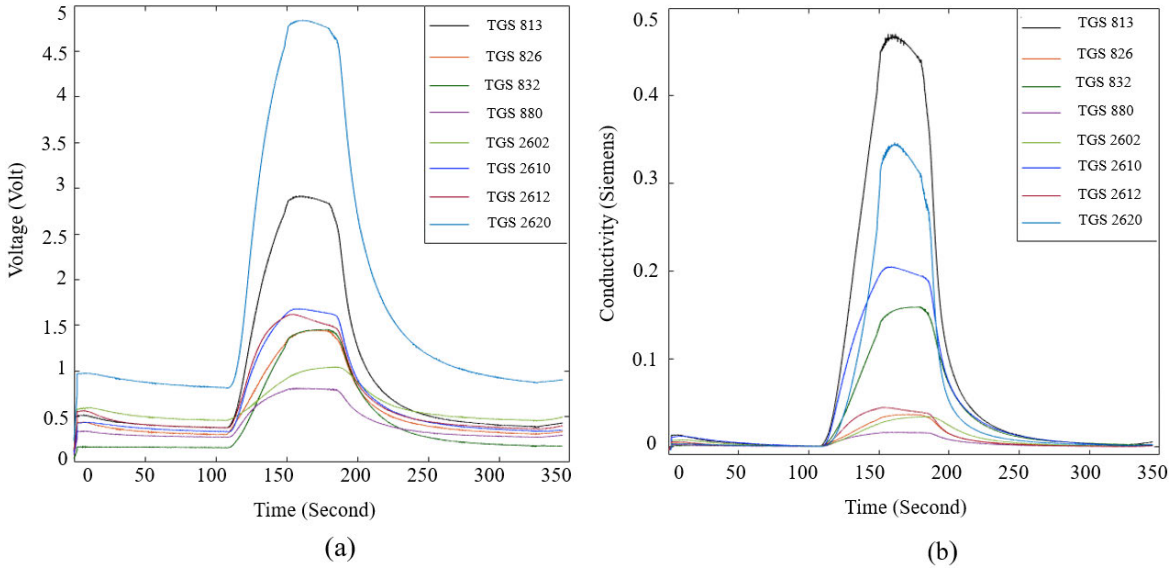

FIGURE 16. The Raw sensor data (left). reference corrected sensor data (right).cancer.

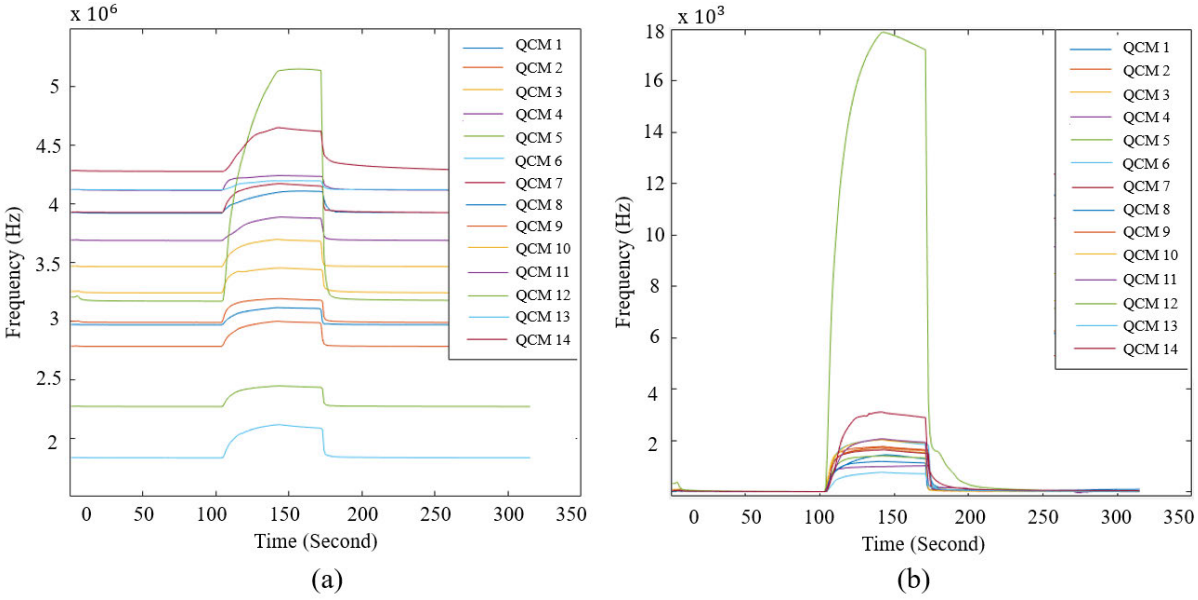

FIGURE 17. The Raw sensor data (a). reference corrected data (b).

$f_{n,s}(t)$ represents the raw frequency data, $f_{n,s}(0)$ represents the resonance frequency value of the sensors, and $f_{n,s,r}(t)$ represent the reference corrected sensor data.

A representative data sample from the reference corrected QCM sensor data and raw QCM sensor data is shown in Figure 17. When the e-nose studies in the literature were examined, it was determined that the conductivity information of the MOS sensor was used rather than the voltage information obtained from the load resistor connected in series with the MOS sensors. In this study, the voltage data obtained from the load resistors connected to the sensors were converted into conductivity data using Equation (3).

$$G_{n,s}(t) = V_{n,s,r}(t) \big/ ((V_c - V_{n,s,r}(t)) R_{L,s} \qquad (3)$$

In equation (3), V_c represents the supply voltage of the sensors, $R_{L,s}$ represents the resistor values connected in series to the MOS sensors, and $G_{n,s}(t))$ represents the conductivity data of the sensors. A representative data sample from the conductivity of the MOS sensor data and reference corrected MOS sensor data is shown in Figure 18. In this study, 338 experiments were performed using the collected breath samples. Of these experiments, 219 were conducted using breath samples from individuals diagnosed with LC, whereas 119 involved breath samples from healthy individuals. Various features were extracted from the data, including maximum, variance, mean, kurtosis, skewness, and gradient over specific temporal intervals. As a result of this process, the size of the feature matrix created using MOS

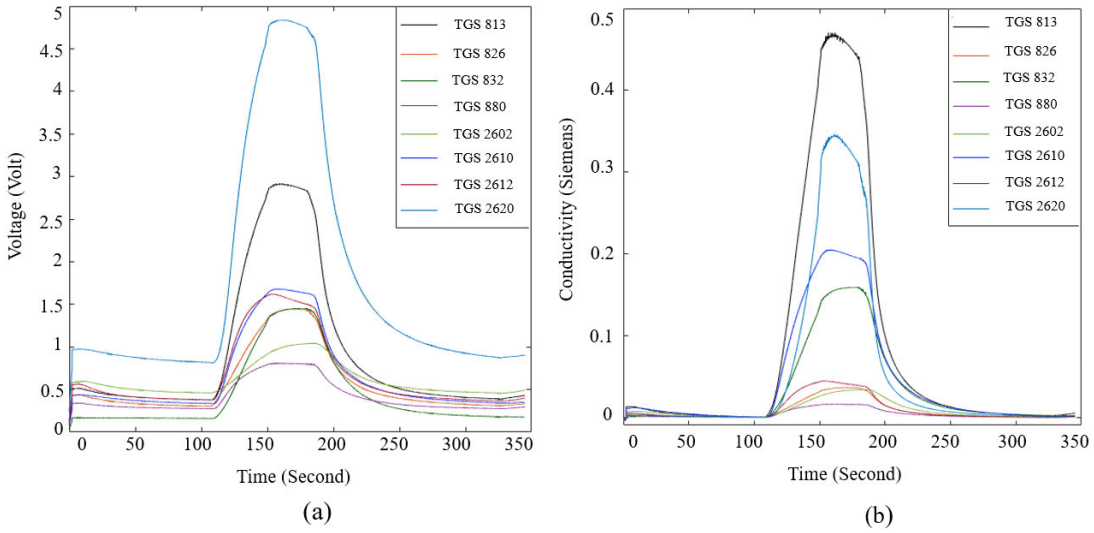

FIGURE 18. The Raw sensor data (a). reference corrected data (b).

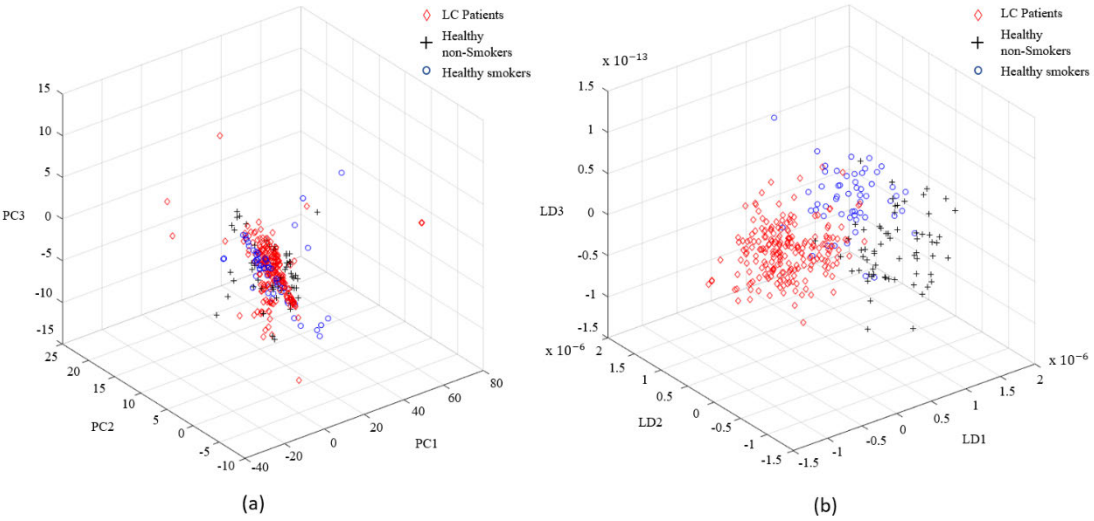

FIGURE 19. The distribution graphs of the first 3 components of the feature matrices obtained by reducing the dimension of the hybrid feature matrix with PCA (a) and LDA (b).

sensor data was (338 × 91), while the size of the feature matrix created using the QCM sensor data was (338 × 7). These two feature matrices are combined to create a hybrid feature matrix of (338 × 98) size. Throughout the remainder of this study, this feature matrix is referred to as the hybrid feature matrix. These feature matrices were classified both in their original sizes and after dimension reduction using linear discriminant analysis (LDA) and principal component analysis (PCA). During the dimension reduction, care was taken to retain at least 90% of the original data. For the MOS sensor data, the initial dimensions of (338×91) were reduced to (338×11) and (338×2) using LDA and PCA, respectively. Similarly, for the QCM sensor data, the initial dimensions of (338×7) were reduced to (338×4) and (338×3) using LDA and PCA, respectively. The size of this hybrid feature matrix was reduced to (338 × 14) using PCA and (338 × 3) using LDA. The distribution graphs of the first three components of the feature matrices obtained by reducing the dimensions of the hybrid feature matrix using PCA and LDA are shown in Figure 19.

In addition to these graphics, the radar plot of the first four features of the hybrid feature matrix, whose dimensions were reduced by LDA, is shown in Figure 20. The shapes in the radar chart are derived from the numerical values in the hybrid feature matrix, the size of which is reduced by LDA. To explain in detail, the shape shown in green in Figure 20(a) is created using the data of the first feature belonging to the patient volunteers. The same method was applied to healthy non-smokers (HnS) and healthy smokers (HS). The shapes shown in yellow and blue in Figure 20(a) were created using the data of the healthy smokers (HS) and healthy non-smokers (HnS) groups in the first feature, respectively. Similarly, Figure 20(b), Figure 20(c), and Figure 20(d) are created using the second, third, and fourth features, respectively.

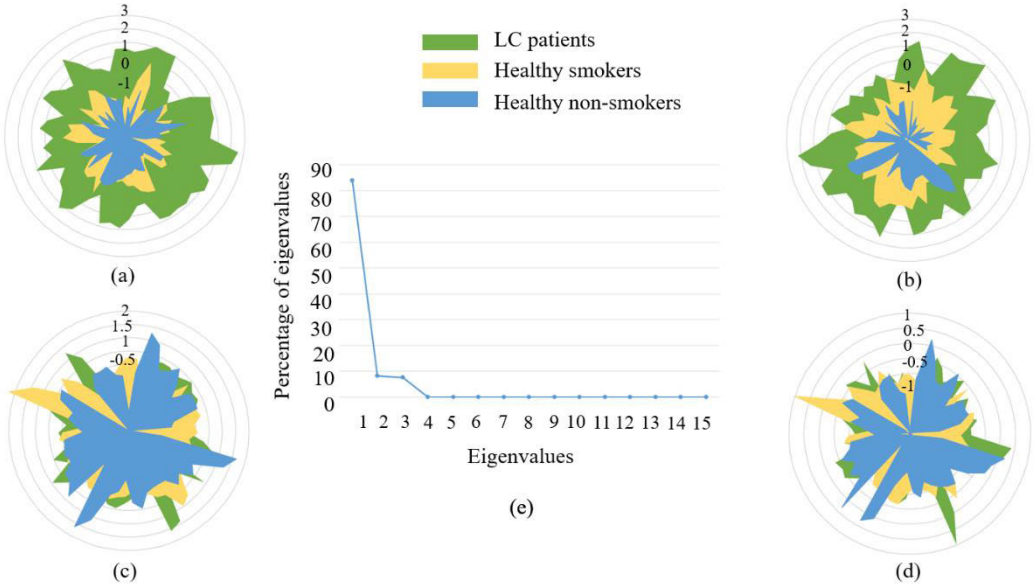

FIGURE 20. The radar plot of the first four features of the hybrid feature matrix whose dimension is reduced with LDA.

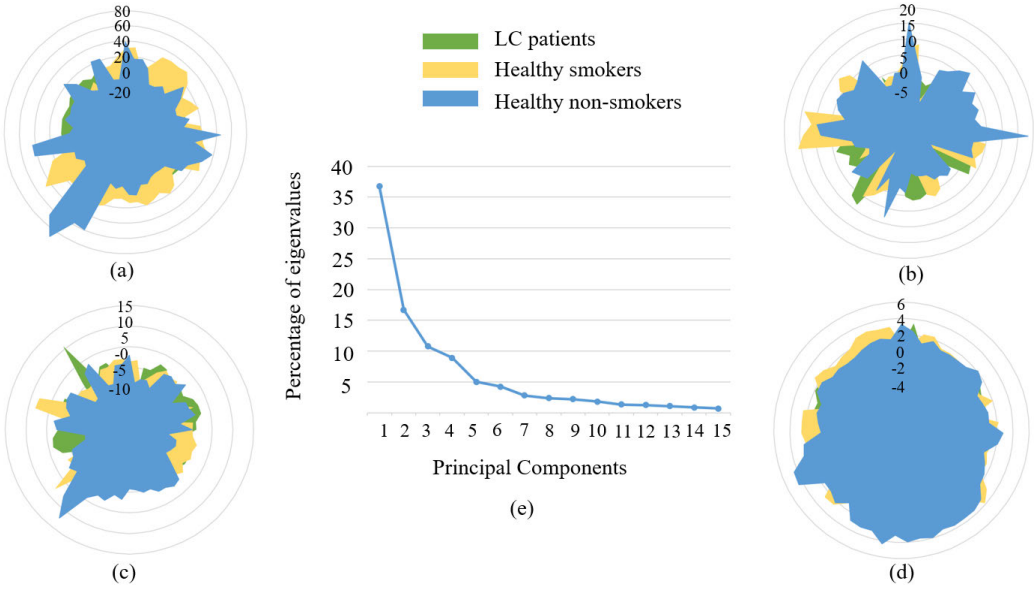

FIGURE 21. The radar plot of the first four features of the hybrid feature matrix whose dimension is reduced with PCA.

Figure 20(e) shows the percentage ratios of the eigenvalues required for the selection of planes to be used in the LDA method. In Figure 20 (e), the first eigenvalue has a very high percentage. This indicates that the first column of the feature matrix obtained after the LDA method has a high discrimination ability. This situation is illustrated in Figure 20. For example, in Figure 20 (a), the shapes representing each class do not completely overlap. However, as shown in Figure 20 (d), the overlap ratio of the shapes increases, and the classes can not be easily distinguished from each other. The radar plot of the first four features of the hybrid feature matrix, whose dimensions were reduced using PCA, is shown in Figure 21.

The shapes in the radar chart are derived from the numerical values in the hybrid feature matrix, the size of which is reduced by PCA. To explain in detail, the shape shown in green in Figure 21(a) is created using the data of the first feature belonging to the patient volunteers. The same method was applied to healthy non-smokers (HnS) and healthy smokers (HS). The shapes shown in yellow and blue in Figure 21(a) were created using the data of the healthy smokers (HS) and healthy non-smokers (HnS) groups in the first feature, respectively. Similarly, Figure 21(b), Figure 21(c), and Figure 21(d) were created using the second, third, and fourth features, respectively. In Figure 21 (e), it can be observed that the percentage value of each feature is small

TABLE 2. The classification results.

Types of features	Classification algorithms					
	DT	L-SVM	Q-SVM	C-SVM	k-NN	RF
MOS	75,34 77,8-72,8-1,90	75,52 77-74,6-1,08	81,12 82,1-79,3-1,17	81,28 82-79,3-1,12	74,3 79,9-71,3-3,30	**81,54** **82,8-79,9-1,09**
PCA(MOS)	85,20 86,1-84,3-0,67	67,86 70,2-64,9-2,17	67.66 70,2-64,9-2,11	85.76 86,4-84,9-0,62	81,06 81,9-79,6-0,90	**87,16** **88,1-86,1-0,88**
LDA(MOS)	88,80 90,1-87,3-1,11	**93.20** **93,1-91,3-0,78**	92,60 92,9-91,7-0,45	91,56 92,8-90,2-1,06	90,10 90,8-89,4-0,51	90,04 90,8-89,1-0,80
QCM	66,82 70,7-64,2-2,70	64,32 64,8-64-0,43	71,96 74,0-70,1-1,46	69,66 73,4-66,5-2,51	70,50 72,2-68,9-1,33	**73,18** **75,1-70,7-1,77**
PCA(QCM)	74.96 76,9-73,7-1,25	67,86 70,2-64,9-2,17	67.66 70,2-64,9-2,11	82,22 84,3-79,3-1,93	**85.96** **87-84,9-0,89**	80,20 81,0-79-0,87
LDA(QCM)	67,30 68,6-65,7-1,30	67,86 70,2-64,9-2,17	67.66 70,2-64,9-2,11	68,92 70,1-68-0,78	66,76 70,4-63,3-2,95	**70,58** **71,2-69,9-0,56**
MOS+QCM	76,06 77,8-74-1,43	76,84 77,8-75,4-0,96	**85,26** **86,4-83,7-1,05**	85,18 86,1-84,1-0,79	75,38 76-74,9-0,48	82,24 82,9-81,4 0,58
PCA (MOS+QCM)	84,24 84,9-83,4-0,60	67,86 70,2-64,9-2,17	67,66 70,2-64,9-2,11	81,46 82,2-80,8-0,63	80,36 85,8-75,4-4,18	**87,56** **88,1-87,2-0,35**
LDA (MOS+QCM)	94,40 94,7-93,8-0,36	94,52 94,7-94,4-0,16	93,80 94,7-92,9-0,73	92.90 94,1-91,6-0,97	92,30 93,8-91,4-0,95	**94,58** **95,6-94,1-0,62**
PCA(MOS)+ PCA(QCM)	83,40 85,2-82,0-1,33	67,86 70,2-64,9-2,17	70,02 70,7-69,2-0,56	83,10 83,7-82,2-0,60	**88,56** **88,8-88,2-0,25**	87,14 88,2-86,10-0,77
LDA(MOS)+ LDA(QCM)	89,80 90,5-88,9-0,65	**93,08** **93,8-92-0,75**	92,60 93,8-92-0,73	90,74 91,4-89,6-0,80	90,60 90,5-89,6-0,35	90,92 91,4-90,2-0,54

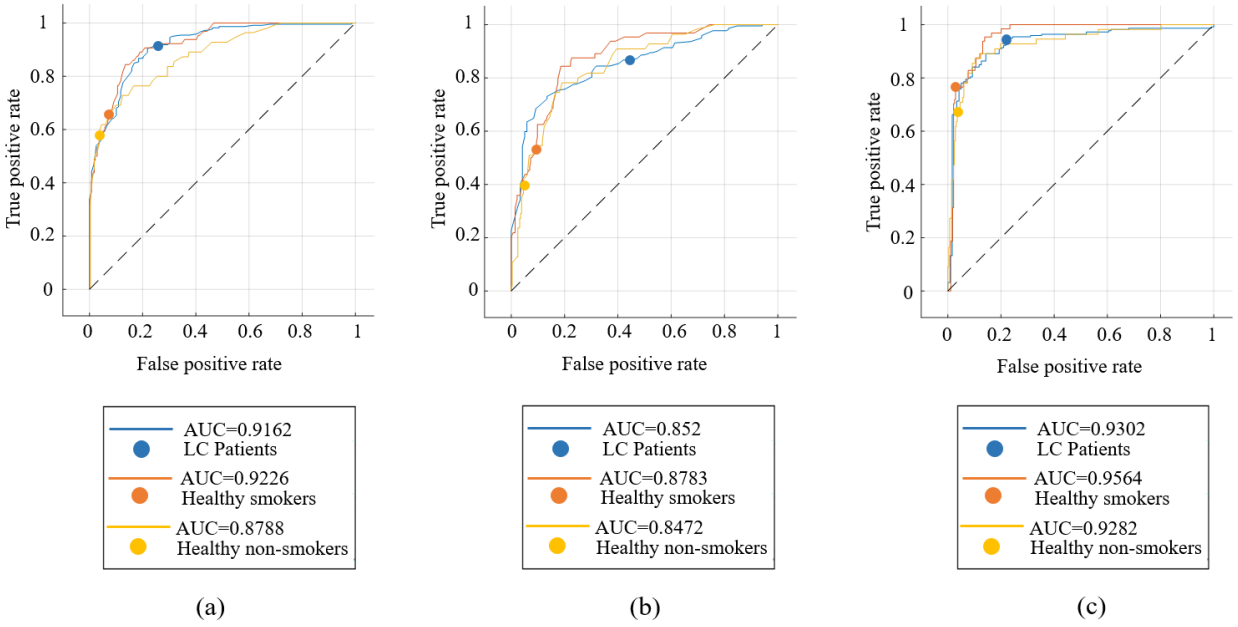

FIGURE 22. Classification details of MOS (a), QCM (b), and MOS+QCM (c) feature matrix.

and close to each other. This means that the feature matrix obtained after PCA does not have a very good ability to distinguish the classes. This situation can also be understood by examining the shapes obtained from the radar plots. In Figure 21, the shapes belonging to each class in the radar plots are not easily separated from each other and overlap significantly. When the classification results in the Results section are examined, it is clear that the feature matrices created using LDA are classified with higher accuracy than those obtained using PCA.

III. DATA CLASSIFICATION AND RESULTS

In this study, data from volunteers categorized as LC, HnS, or HS were classified into three distinct groups using various classification algorithms. The constructed feature matrices were classified using decision tree (DT), random forest (RF), k-nearest neighbor (k-NN), and support vector machine (SVM) algorithms, both in their original forms and after dimension reduction using PCA and LDA methods. The classification results are presented in Table 2. In this table, the feature matrices are labeled using abbreviations that indicate the sensor data features and the applied dimensionality reduction method. For example, the features of the matrix derived from the MOS sensors data without dimension reduction are denoted as MOS (in the rest of the study, the names of the feature matrices are underlined), the matrix reduced using the PCA method is referred to as PCA(MOS), and the matrix reduced using the LDA method is denoted as

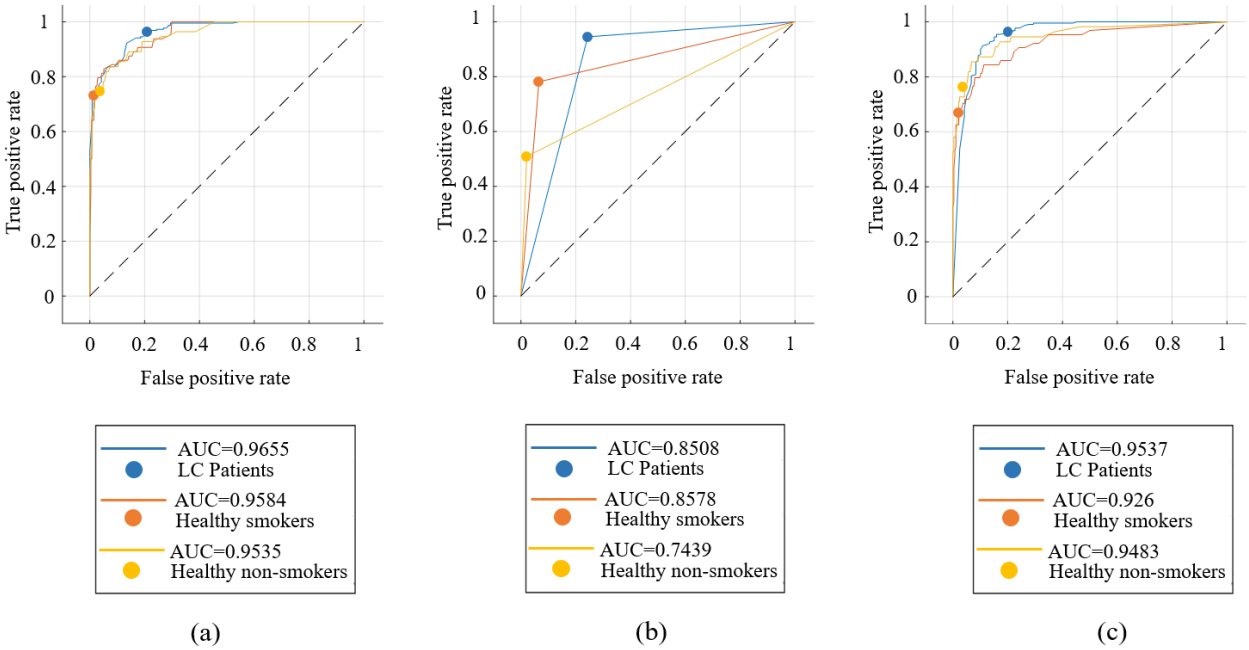

FIGURE 23. Classification details of PCA(MOS) (a), PCA(QCM) (b), and PCA(MOS+QCM) (c) feature matrix.

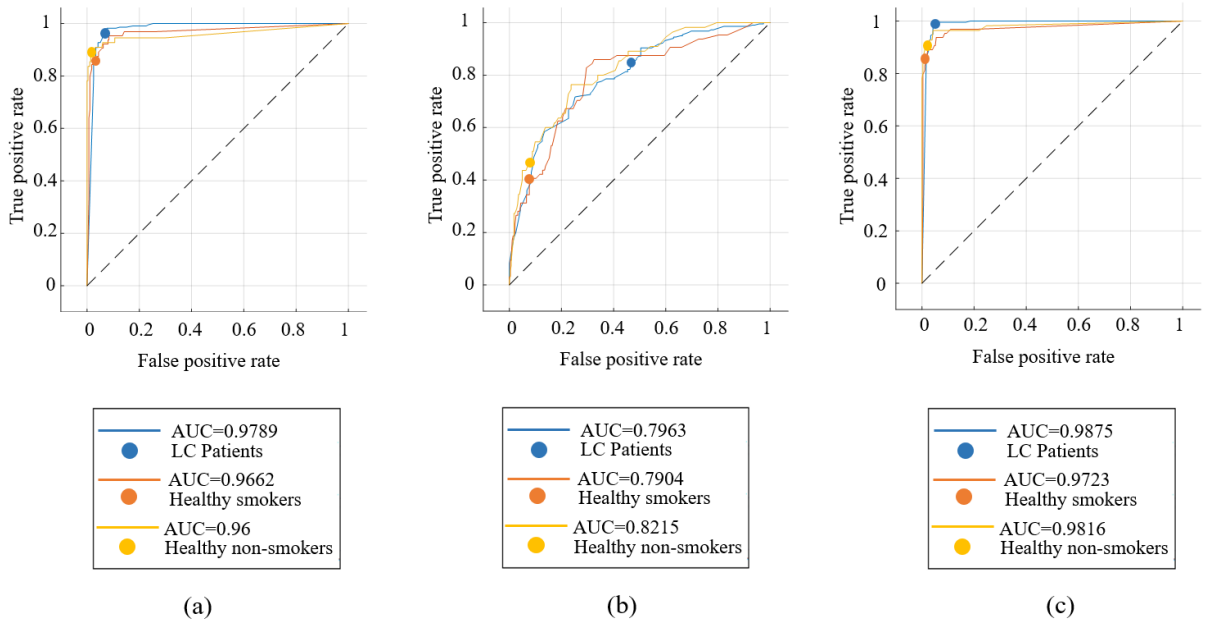

FIGURE 24. Classification details of LDA(MOS) (a), LDA(QCM) (b), and LDA(MOS+QCM) (c) feature matrix.

LDA(MOS). A 5-fold cross-validation technique was used to evaluate the success of classifiers, and each classification algorithm was executed ten times for each feature matrix. The results include the average accuracy, highest accuracy, lowest accuracy, and standard deviation values. In addition to the results summarized in Table 2, receiver operating characteristic (ROC) curves were generated to provide further insight into the classification outcomes. Initially, ROC curves were created for the classification of the feature matrices without dimension reduction. These curves, presented in Figure 22, illustrate the classification results for the MOS, QCM, and MOS+QCM feature matrices. An analysis of Figure 22 highlights the significant impact of sensor type on classification performance, with the MOS feature matrix achieving a higher classification accuracy than the QCM matrix. One of the primary goals of this study was to enhance classification accuracy by integrating data from different sensor types. As shown in Figure 22, this objective was achieved because the MOS+QCM feature matrix outperformed the MOS and QCM matrices when classified

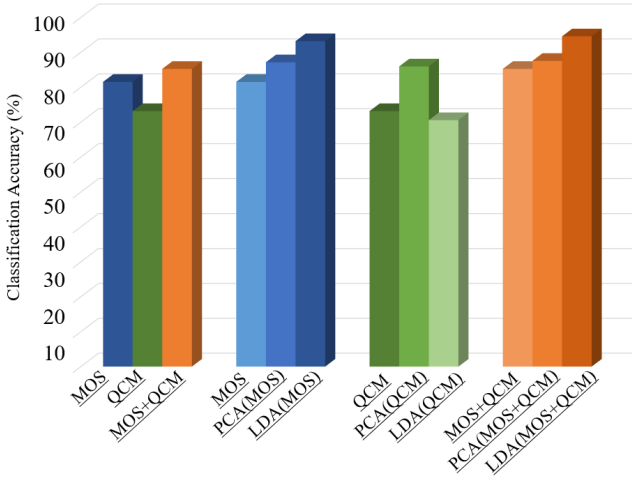

FIGURE 25. A comparative graph of classification results.

separately. To further assess the impact of the PCA and LDA algorithms on classification performance, ROC curves were generated for the dimensionally reduced data. Figures 23 and 24 show the ROC curves of the feature matrix.

Dimensions were reduced using PCA and LDA, respectively. These results indicate that, even after dimension reduction, the combined MOS+QCM feature matrix consistently delivers the highest classification accuracy. Additionally, a comparative analysis of Figures 21 and 22 reveals that the LDA algorithm has a more pronounced positive effect on the classification performance than the PCA algorithm. In figures 20, 21, and 22, the AUC represents the area under the curve.

IV. CONCLUSION

In this study, a hybrid sensor-based e-nose circuit was developed by incorporating 8 MOS sensors and 14 QCM sensors. When studies on LC detection using e-noses in the literature were examined, it was observed that the majority of e-noses contained a single type of gas sensor. This includes commercial e-noses and those developed by other researchers. Experiments were conducted using breath samples collected from 60 patients with LC, 20 patients with HnS, and 20 patients with HS. The experimental data were analyzed to identify features that could effectively differentiate the breath samples of LC, HnS, and HS volunteers. The resulting feature matrices were classified using various classification algorithms, including the decision tree (DT), k-nearest neighbor (k-NN), support vector machine (SVM), and random forest (RF). When the MOS and QCM feature matrices were classified individually, the highest classification accuracies were 81.54% and 73.18%, respectively. Subsequently, a hybrid feature matrix (MOS + QCM) was created by combining the MOS and QCM feature matrices, thereby achieving an improved classification accuracy of 85.26%. These results support the initial hypothesis of this study, which proposed that combining data from different sensor types could enhance the LC detection accuracy. In the initial phase, classification was performed on the original feature matrices without dimension reduction. To further investigate the impact of dimension reduction on classification performance, LDA and PCA algorithms were applied to reduce the dimensions of the feature matrices, followed by classification of the dimension-reduced feature matrices.

Based on the classification process performed using the MOS feature matrix, the highest classification accuracy achieved was 81.54%. In the classification operations performed using PCA(MOS) and LDA(MOS) feature matrices, the classification accuracy increased to 87.16% and 93.20%, respectively. The highest classification accuracy obtained from the classification performed using the QCM feature matrix was 73.18%. In the classification process performed using PCA(QCM) and LDA(QCM) feature matrices, the classification accuracy values were 85.96% and 70.58%, respectively. In the classification process performed using the hybrid feature matrix (MOS+QCM), the highest classification accuracy was 85.26%. In the classification operations performed using PCA(MOS+QCM) and LDA(MOS+QCM) feature matrices, the classification accuracy values obtained were 87.56% and 94.58%, respectively. A comparative graph showing the effects of both the feature matrix types used and the dimensionality reduction algorithms on the classification accuracy is shown in Figure 25. Figure 25 clearly demonstrates that leveraging data from different types of gas sensors can enhance the performance of the system. Thus, an inexpensive electronic nose that can detect LC with high accuracy using breath analysis has been developed. All procedures and results carried out during this study bring this work to the forefront by distinguishing it from the examples found in the literature. Based on these results, we anticipate that our study will have broad practical applications for LC detection.

REFERENCES

[1] R. L. Siegel, K. D. Miller, A. G. Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, "Colorectal cancer statistics, 2020," *CA, Cancer J. Clinicians*, vol. 70, no. 3, pp. 145–164, Mar. 2020.

[2] M. D. Luque de Castro and M. A. Fernández-Peralbo, "Analytical methods based on exhaled breath for early detection of lung cancer," *TrAC Trends Anal. Chem.*, vol. 38, pp. 13–20, Sep. 2012.

[3] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 71, no. 3, pp. 209–249, Feb. 2021.

[4] The National Lung Screening Trial Research Team, "Reduced lung-cancer mortality with low-dose computed tomographic screening," *New England J. Med.*, vol. 365, no. 5, pp. 395–409, Aug. 2011.

[5] S. Blandin Knight, P. A. Crosbie, H. Balata, J. Chudziak, T. Hussell, and C. Dive, "Progress and prospects of early detection in lung cancer," *Open Biol.*, vol. 7, no. 9, Sep. 2017, Art. no. 170070.

[6] P. M. Marcus, E. J. Bergstralh, M. H. Zweig, A. Harris, K. P. Offord, and R. S. Fontana, "Extended lung cancer incidence follow-up in the mayo lung project and overdiagnosis," *J. Nat. Cancer Inst.*, vol. 98, no. 11, pp. 748–756, Jun. 2006.

[7] L. V. Bel'skaya, E. A. Sarf, V. K. Kosenok, and I. A. Gundyrev, "Biochemical markers of saliva in lung cancer: Diagnostic and prognostic perspectives," *Diagnostics*, vol. 10, no. 4, p. 186, Mar. 2020.

[8] P. A. VanderLaan, S. Roy-Chowdhuri, C. C. Griffith, V. L. Weiss, and C. N. Booth, "Molecular testing of cytology specimens: Overview of assay selection with focus on lung, salivary gland, and thyroid testing," *J. Amer. Soc. Cytopathol.*, vol. 11, no. 6, pp. 403–414, Nov. 2022.

[9] S. Herath, H. Sadeghi Rad, P. Radfar, R. Ladwa, M. Warkiani, K. O'Byrne, and A. Kulasinghe, "The role of circulating biomarkers in lung cancer," *Frontiers Oncol.*, vol. 11, Jan. 2022, Art. no. 801269.

[10] S. Wang, K. Zhang, S. Tan, J. Xin, Q. Yuan, H. Xu, X. Xu, Q. Liang, D. C. Christiani, M. Wang, L. Liu, and M. Du, "Circular RNAs in body fluids as cancer biomarkers: The new frontier of liquid biopsies," *Mol. Cancer*, vol. 20, no. 1, pp. 1–10, Jan. 2021.

[11] M. Balbisi, S. Sugár, and L. Turiák, "Protein glycosylation in lung cancer from a mass spectrometry perspective," *Mass Spectrometry Rev.*, vol. 2024, pp. 1–21, Apr. 2024.

[12] S.-R. Yang, A. M. Schultheis, H. Yu, D. Mandelker, M. Ladanyi, and R. Büttner, "Precision medicine in non-small cell lung cancer: Current applications and future directions," *Seminars Cancer Biol.*, vol. 84, pp. 184–198, Sep. 2022.

[13] F. Passiglia, M. Cinquini, L. Bertolaccini, M. Del Re, F. Facchinetti, R. Ferrara, T. Franchina, A. R. Larici, U. Malapelle, J. Menis, A. Passaro, S. Pilotto, S. Ramella, G. Rossi, R. Trisolini, and S. Novello, "Benefits and harms of lung cancer screening by chest computed tomography: A systematic review and meta-analysis," *J. Clin. Oncol.*, vol. 39, no. 23, pp. 2574–2585, Aug. 2021.

[14] M. T. Lu, V. K. Raghu, T. Mayrhofer, H. J. W. L. Aerts, and U. Hoffmann, "Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: Development and validation of a prediction model," *Ann. Internal Med.*, vol. 173, no. 9, pp. 704–713, Aug. 2020.

[15] L. L. Cheng, Z. Kobus, M. Bobus, L. Su, and D. C. Christiani, "Investigations of serum metabolomic relationships between lung cancer and Alzheimer's disease using nuclear magnetic resonance spectroscopy," *Alzheimer's Dementia*, vol. 19, Dec. 2023, Art. no. e082256.

[16] Y. Wei, C. Yang, H. Jiang, Q. Li, F. Che, S. Wan, S. Yao, F. Gao, T. Zhang, J. Wang, and B. Song, "Multi-nuclear magnetic resonance spectroscopy: State of the art and future directions," *Insights Imag.*, vol. 13, no. 1, p. 135, Aug. 2022.

[17] H. L. Lancaster, M. A. Heuvelmans, and M. Oudkerk, "Low-dose computed tomography lung cancer screening: Clinical evidence and implementation research," *J. Internal Med.*, vol. 292, no. 1, pp. 68–80, Mar. 2022.

[18] S. A. Fedewa, E. A. Kazerooni, J. L. Studts, R. A. Smith, P. Bandi, A. G. Sauer, M. Cotter, H. M. Sineshaw, A. Jemal, and G. A. Silvestri, "State variation in low-dose computed tomography scanning for lung cancer screening in the United States," *J. Nat. Cancer Inst.*, vol. 113, no. 8, pp. 1044–1052, Aug. 2021.

[19] MM Oken, WG Hocking, PA Kvale, GL Andriole "Screening by chest radiograph and lung cancer mortality," *JAMA*, vol. 306, no. 17, p. 1865, Oct. 2011.

[20] W. Li, H.-Y. Liu, Z.-R. Jia, P.-P. Qiao, X.-T. Pi, J. Chen, and L.-H. Deng, "Advances in the early detection of lung cancer using analysis of volatile organic compounds: From imaging to sensors," *Asian Pacific J. Cancer Prevention*, vol. 15, no. 11, pp. 4377–4384, Jun. 2014.

[21] P. J. Mazzone, N. Obuchowski, M. Phillips, B. Risius, B. Bazerbashi, and M. Meziane, "Lung cancer screening with computer aided detection chest radiography: Design and results of a randomized, controlled trial," *PLoS ONE*, vol. 8, no. 3, Mar. 2013, Art. no. e59650.

[22] E. Jantus-Lewintre, M. Usó, E. Sanmartin, and C. Camps, "Update on biomarkers for the detection of lung cancer," *Lung Cancer, Targets Therapy*, vol. 2012, pp. 21–29, Jun. 2012.

[23] W. Miekisch, J. K. Schubert, and G. F. E. Noeldge-Schomburg, "Diagnostic potential of breath analysis–focus on volatile organic compounds," *Clinica Chim. Acta*, vol. 347, nos. 1–2, pp. 25–39, Sep. 2004.

[24] A. Amann, B. D. L. Costello, W. Miekisch, J. Schubert, B. Buszewski, J. Pleil, N. Ratcliffe, and T. Risby, "The human volatilome: Volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva," *J. Breath Res.*, vol. 8, no. 3, Jun. 2014, Art. no. 034001.

[25] T. Issitt, L. Wiggins, M. Veysey, S. T. Sweeney, W. J. Brackenbury, and K. Redeker, "Volatile compounds in human breath: Critical review and meta-analysis," *J. Breath Res.*, vol. 16, no. 2, Feb. 2022, Art. no. 024001.

[26] A. Sharma, R. Kumar, and P. Varadwaj, "Smelling the disease: Diagnostic potential of breath analysis," *Mol. Diagnosis Therapy*, vol. 27, no. 3, pp. 321–347, Feb. 2023.

[27] P. D. Wagner, "Blood gas transport: Carriage of oxygen and carbon dioxide in blood," *Seminars Respiratory Crit. Care Med.*, vol. 44, no. 5, pp. 569–583, Aug. 2023.

[28] L. Pauling, A. B. Robinson, R. Teranishi, and P. Cary, "Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography," *Proc. Nat. Acad. Sci. USA*, vol. 68, no. 10, pp. 2374–2376, Oct. 1971.

[29] Y. Yang, T. Cui, D. Li, S. Ji, Z. Chen, W. Shao, H. Liu, and T.-L. Ren, "Breathable electronic skins for daily physiological signal monitoring," *Nano-Micro Lett.*, vol. 14, no. 1, p. 161, Aug. 2022.

[30] B. Behera, R. Joshi, G. K. Anil Vishnu, S. Bhalerao, and H. J. Pandya, "Electronic nose: A non-invasive technology for breath analysis of diabetes and lung cancer patients," *J. Breath Res.*, vol. 13, no. 2, Mar. 2019, Art. no. 024001.

[31] N. Ratcliffe, T. Wieczorek, N. Drabińska, O. Gould, A. Osborne, and B. De Lacy Costello, "A mechanistic study and review of volatile products from peroxidation of unsaturated fatty acids: An aid to understanding the origins of volatile organic compounds from the human body," *J. Breath Res.*, vol. 14, no. 3, May 2020, Art. no. 034001.

[32] A. S. Oyerinde, V. Selvaraju, J. R. Babu, and T. Geetha, "Potential role of oxidative stress in the production of volatile organic compounds in obesity," *Antioxidants*, vol. 12, no. 1, p. 129, Jan. 2023.

[33] N. Drabińska, C. Flynn, N. Ratcliffe, I. Belluomo, A. Myridakis, O. Gould, M. Fois, A. Smart, T. Devine, and B. D. L. Costello, "A literature survey of all volatiles from healthy human breath and bodily fluids: The human volatilome," *J. Breath Res.*, vol. 15, no. 3, Apr. 2021, Art. no. 034001.

[34] J. O. Ogbodo, A. V. Arazu, T. C. Iguh, N. J. Onwodi, and T. C. Ezike, "Volatile organic compounds: A proinflammatory activator in autoimmune diseases," *Frontiers Immunol.*, vol. 13, Jul. 2022, Art. no. 928379.

[35] M. Karunagaran, P. Ramani, S. Gheena, R. Abilasha, and R. Hannah, "Volatile organic compounds in human breath," *Indian J. Dental Res.*, vol. 33, no. 1, pp. 100–104, Jan. 2022.

[36] J. Kokoszka, R. L. Nelson, W. I. Swedler, J. Skosey, and H. Abcarian, "Determination of inflammatory bowel disease activity by breath pentane analysis," *Diseases Colon Rectum*, vol. 36, no. 6, pp. 597–601, Jun. 1993.

[37] J. Scholpp, J. K. Schubert, W. Miekisch, and K. Geiger, "Breath markers and soluble lipid peroxidation markers in critically ill patients," *Clin. Chem. Lab. Med.*, vol. 40, no. 6, pp. 587–594, Jan. 2002.

[38] H. M. Saraoglu, A. O. Selvi, M. A. Ebeoglu, and C. Tasaltin, "Electronic nose system based on quartz crystal microbalance sensor for blood glucose and HbA1c levels from exhaled breath odor," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4229–4235, Nov. 2013.

[39] H. Huang, J. Zhou, S. Chen, L. Zeng, and Y. Huang, "A highly sensitive QCM sensor coated with Ag+-ZSM-5 film for medical diagnosis," *Sens. Actuators B, Chem.*, vol. 101, no. 3, pp. 316–321, Jul. 2004.

[40] K. Anastassakis, *Androgenetic Alopecia from A To Z*. Cham, Swizerland: Springer, 2022.

[41] C.-J. Chen, M.-C. Cheng, C.-N. Hsu, and Y.-L. Tain, "Sulfur-containing amino acids, hydrogen sulfide, and sulfur compounds on kidney health and disease," *Metabolites*, vol. 13, no. 6, p. 688, May 2023.

[42] S. Davies, P. Spanel, and D. Smith, "Quantitative analysis of ammonia on the breath of patients in end-stage renal failure," *Kidney Int.*, vol. 52, no. 1, pp. 223–228, Jul. 1997.

[43] G. Peng, M. Hakim, Y. Y. Broza, S. Billan, R. Abdah-Bortnyak, A. Kuten, U. Tisch, and H. Haick, "Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors," *Brit. J. Cancer*, vol. 103, no. 4, pp. 542–551, Jul. 2010.

[44] D. Poli, M. Goldoni, M. Corradi, O. Acampa, P. Carbognani, E. Internullo, A. Casalini, and A. Mutti, "Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME–GC/MS," *J. Chromatography B*, vol. 878, no. 27, pp. 2643–2651, Oct. 2010.

[45] Z. Jia, A. Patra, V. K. Kutty, and T. Venkatesan, "Critical review of volatile organic compound analysis in breath and in vitro cell culture for detection of lung cancer," *Metabolites*, vol. 9, no. 3, p. 52, Mar. 2019.

[46] R. Su, T. Yang, X. Zhang, N. Li, X. Zhai, and H. Chen, "Mass spectrometry for breath analysis," *TrAC Trends Anal. Chem.*, vol. 158, Jan. 2022, Art. no. 116823.

[47] Y.-Y. Zhou, J.-P. Chen, L. Gan, W. Xu, Y. Liu, Y.-G. Zhao, and Y. Zhu, "A non-invasive method for the detection of glucose in human exhaled breath by condensation collection coupled with ion chromatography," *J. Chromatography A*, vol. 1685, Dec. 2022, Art. no. 463564.

[48] T. Seesaard, N. Goel, M. Kumar, and C. Wongchoosuk, "Advances in gas sensors and electronic nose technologies for agricultural cycle applications," *Comput. Electron. Agricult.*, vol. 193, Feb. 2022, Art. no. 106673.
[49] E. Kress-Rogers, *Handbook of Biosensors and Electronic Noses*. Boca Raton, FL, USA: CRC Press, 2024.
[50] S. Dhall, B. R. Mehta, A. K. Tyagi, and K. Sood, "A review on environmental gas sensors: Materials and technologies," *Sensors Int.*, vol. 2, Jan. 2021, Art. no. 100116.
[51] S. Feng, F. Farha, Q. Li, Y. Wan, Y. Xu, T. Zhang, and H. Ning, "Review on smart gas sensing technology," *Sensors*, vol. 19, no. 17, p. 3760, Aug. 2019.
[52] M. Rivai and D. Aulia, "Use of electronic nose to identify levels of cooking cookies," *IEEE Access*, vol. 12, pp. 97235–97247, 2024.
[53] A. Khorramifar, H. Karami, L. Lvova, A. Kolouri, E. Łazuka, M. Piłat-Rożek, G. Łagód, J. Ramos, J. Lozano, M. Kaveh, and Y. Darvishi, "Environmental engineering applications of electronic nose systems based on MOX gas sensors," *Sensors*, vol. 23, no. 12, p. 5716, Jun. 2023.
[54] M. Abdelkhalek, S. Alfayad, F. Benouezdou, M. B. Fayek, and L. Chassagne, "Compact and embedded electronic nose for volatile and non-volatile odor classification for robot applications," *IEEE Access*, vol. 7, pp. 98267–98276, 2019.
[55] Y. Li, Z. Wang, T. Zhao, H. Li, J. Jiang, and J. Ye, "Electronic nose for the detection and discrimination of volatile organic compounds: Application, challenges, and perspectives," *TrAC Trends Anal. Chem.*, vol. 180, Nov. 2024, Art. no. 117958.
[56] W. S. Al-Dayyeni, S. Al-Yousif, M. M. Taher, A. W. Al-Faouri, N. M. Tahir, M. M. Jaber, F. Ghabban, I. A. Najm, I. M. Alfadli, O. Z. Ameerbakhsh, M. J. Mnati, N. A. Al-Shareefi, and A. H. Saleh, "A review on electronic nose: Coherent taxonomy, classification, motivations, challenges, recommendations and datasets," *IEEE Access*, vol. 9, pp. 88535–88551, 2021.
[57] M. Aleixandre and T. Nakamoto, "Online learning for active odor sensing based on a QCM gas sensor array and an odor blender," *IEEE Sensors J.*, vol. 22, no. 23, pp. 22310–22318, Dec. 2022.
[58] S. Okur, M. Sarheed, R. Huber, Z. Zhang, L. Heinke, A. Kanbar, C. Wöll, P. Nick, and U. Lemmer, "Identification of mint scents using a QCM based E-nose," *Chemosensors*, vol. 9, no. 2, p. 31, Feb. 2021.
[59] Malikhah, R. Sarno, S. Inoue, M. S. H. Ardani, D. P. Purbawa, S. I. Sabilla, K. R. Sungkono, C. Fatichah, D. Sunaryono, A. Bakhtiar, Libriansyah, C. R. S. Prakoeswa, D. Tinduh, and Y. Hernaningsih, "Detection of infectious respiratory disease through sweat from axillary using an E-nose with stacked deep neural network," *IEEE Access*, vol. 10, pp. 51285–51298, 2022.
[60] D. Aulia, R. Sarno, S. C. Hidayati, and M. Rivai, "Optimization of the electronic nose sensor array for asthma detection based on genetic algorithm," *IEEE Access*, vol. 11, pp. 74924–74935, 2023.
[61] R. Gasparri, M. Santonico, C. Valentini, G. Sedda, A. Borri, F. Petrella, P. Maisonneuve, G. Pennazza, A. D'Amico, C. Di Natale, R. Paolesse, and L. Spaggiari, "Volatile signature for the early diagnosis of lung cancer," *J. Breath Res.*, vol. 10, no. 1, Feb. 2016, Art. no. 016007.
[62] R. Blatt, A. Bonarini, E. Calabro, M. D. Torre, M. Matteucci, and U. Pastorino, "Lung cancer identification by an electronic nose based on an array of MOS sensors," in *Proc. Int. Joint Conf. Neural Netw.*, Aug. 2007, pp. 1423–1428.
[63] M. K. Nakhleh et al., "Diagnosis and classification of 17 diseases from 1404 subjects via pattern analysis of exhaled molecules," *ACS Nano*, vol. 11, no. 1, pp. 112–125, Dec. 2016.
[64] V. A. Binson, M. Subramoniam, and L. Mathew, "Discrimination of COPD and lung cancer from controls through breath analysis using a self-developed e-nose," *J. Breath Res.*, vol. 15, no. 4, Oct. 2021, Art. no. 046003.
[65] R. van de Goor, M. van Hooren, A.-M. Dingemans, B. Kremer, and K. Kross, "Training and validating a portable electronic nose for lung cancer screening," *J. Thoracic Oncol.*, vol. 13, no. 5, pp. 676–681, May 2018.
[66] T. Saidi, M. Moufid, K. De Jesus Beleño-Saenz, T. G. Welearegay, N. E. Bari, A. Lisset Jaimes-Mogollon, R. Ionescu, J. E. Bourkadi, J. Benamor, M. E. Ftouh, and B. Bouchikhi, "Non-invasive prediction of lung cancer histological types through exhaled breath analysis by UV-irradiated electronic nose and GC/QTOF/MS," *Sens. Actuators B, Chem.*, vol. 311, May 2020, Art. no. 127932.
[67] M. H. M. C. Scheepers, Z. Al-Difaie, L. Brandts, A. Peeters, B. van Grinsven, and N. D. Bouvy, "Diagnostic performance of electronic noses in cancer diagnoses using exhaled breath," *JAMA Netw. Open*, vol. 5, no. 6, Jun. 2022, Art. no. e2219372.

UMIT OZSANDIKCIOGLU received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Karadeniz Technical University, Trabzon, Türkiye, in 2013, 2016, and 2023, respectively. He is currently an Academician with the Central Research Laboratory Application and Research Center, Karadeniz Technical University. His research interests include system analysis and design, signal processing, pattern recognition, and electronic noses.

AYTEN ATASOY received the B.S., M.S., and Ph. D. degrees in electrical and electronics engineering from Karadeniz Technical University, Trabzon, Türkiye, in 1987, 1992, and 1997, respectively. She is currently a Professor with the Department of Electrical and Electronics Engineering, Karadeniz Technical University. Her research interests include system analysis and design, system identification and estimation, signal processing, pattern recognition, and electronic noses.

YUSUF SEVIM received the B.S., M.S., and Ph.D. degrees in electrical and electronics engineering from Karadeniz Technical University, Trabzon, Türkiye, in 2000, 2003, and 2009, respectively. He is currently an Associate Professor Doctor with the Department of Electrical and Electronics Engineering, Karadeniz Technical University. His research interests include system analysis and design, signal processing, and pattern recognition.