



Analysis Space Transformation Based Electronic Nose for Efficient Detection and Monitoring of Volatile Organic Compounds, Gases/Odors in Smart Homes

Kanak Kumar¹ , Navin Singh Rajput^{1*} 

¹Department of Electronics Engineering, Indian Institute of Technology (Banaras Hindu University), Varanasi - 221005, UP, India

E-mail: {kanakkumar.rs.ece18, nsrajput.ece}@iitbhu.ac.in

Corresponding Authors *

Abstract

In this work, we have developed an efficient electronic nose (E-nose) for a smart home to detect and monitor various volatile organic compounds (VOCs) that are usually generated in the household sector during several daily chores (e.g., worship, safety & security, hygiene, LPG leakage detection, cooking and smoking). We have implemented the proposed electronic nose using a six-element low-cost MQ-series-based gas sensor array and aESP-32 Microcontroller. The raw gas sensor array responses are first transformed in the Standardised Principal Component Analysis (SPCA) based analysis domain. In this domain, the gas sensor data shows well shaped and separated clusters belonging to various VOCs. A simplistic classifier is then designed using Artificial Neural Network (ANN) trained in the SPCA transformed domain itself which outperforms in classification accuracy. In this experiment, we have captured 2465 training samples belonging to 17 diverse VOCs (odors and smokes) commonly found in household ambience. Another 85 samples were also captured for testing purposes which were not used during the training of the ANN classifier. The accuracy in correct classification has been 96.47% for the 17-class test samples while the precision, recall and F1-score were 96%, 95% and 95%, respectively. The mean squared error (MSE) in this experiment was between 1.35×10^{-6} and 2.15×10^{-2} with an average MSE of 1.42×10^{-3} .

Keywords: Gas Sensor System, Principal Component Analysis, ANN Classifier, Electronic Nose, Gas Sensors, ESP 32 Microcontroller

1. Introduction

Electronic noses (E-noses) have been found useful in healthcare [1], Agriculture Monitoring [2], Biometric Applications and Food Industry [3], Security [4], and Safety Applications [5] and their performance enhancement is a never-ending research work. E-noses are integrated with Pattern Recognition (PR) techniques for identification and estimation of various volatile organic compounds (VOCs), present in the ambience. E-noses help us to prevent unexpected financial, health/life, and other losses. These can also detect and identify various inflammable, non-inflammable, toxic, and hazardous gases/odors. Further, thick-film gas sensor arrays are based on tin-oxide materials duly doped with wide variety of dopants. These are affordable and effective for detecting various gases and VOCs without much poisoning [5]. A basic block schematic of an E-nose is depicted below in Figure 1.

In recent literature, Artificial Neural Network (ANN) classifiers have been used to classify certain gases/odors. In contrast, Rajput et al. (2010) have used the standardized principal component analysis (SPCA) domain to classify and quantify four different types of gases/odors, viz., acetone, carbon tetrachloride, ethyl methyl ketone and xylene. They have used a two-stage neural network system for performance enhancement [6]. Xu et al. (2022) and Olenevaa et al. (2019) have used ANN, Support Vector Machines (SVM) with Radial Basis Function (RBF) kernel, and Random Forest (RF) classification approaches [7], [8]. Rabeb et al. (2015) have described a laboratory-designed WO₃ gas sensor array to determine the fingerprints of gas samples. Ozone, ethanol, acetone, and a binary mixture of ozone and ethanol have been exposed on the

sensor array, at a constant temperature. Through PCA and SVM, they achieved well classification accuracy [9]. Andrew (2016) used a highly complicated hybrid model with the PCA-Probabilistic Neural Network (PNN) method to diagnose incipient-stage fires in buildings using 1000 samples and three principal components [10]. Xu et al. (2018) developed a hybrid gas identification, and concentration measurement system using K-Nearest Neighbour (KNN) classification, after modelling nonlinear characteristics of binary mixtures of CO and CH₄ gas components using KPCA [11].

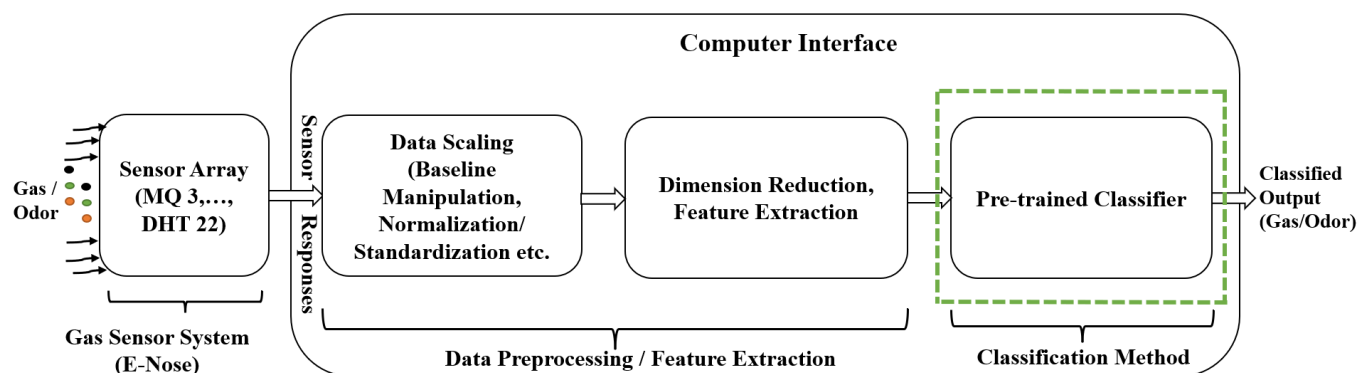


Figure 1. Basic Block Schematic of an E-nose

Furthermore, Zhao et al. (2016) developed a four-sensor array using neural networks to differentiate formaldehyde from three interfering VOCs (C₃H₆O, C₇H₈ and C₂H₅OH) using 108 gas samples. This work used the different variants of neural networks: the ELM with PCA and the BPNN with SVM. The PCA helps to increase the ELM's accuracy by pre-processing the sensor data, while the SVM approach yields the best accuracy [12]. Olenevaa et al. (2019) developed a quick and straightforward analytical technique to identify children's toys tainted with possibly harmful VOCs. They have used QCM based gas sensor array of eight sensor elements, each optimized for different sorbents, to classify toys into clean and dangerous ones with an accuracy of about 96%. This sensor array has been quite costly and uses complex versions of Linear Discriminant Analysis (LDA) and SVM classification models [8]. Unfortunately, they take much time and are disruptive when measuring samples. These devices are also very costly and enormous. Usually, literature represents only the study of various gases/odors produced in a particular paradigm. For example, in [13], the authors have used an e-nose for monitoring the microbial spoilage of canned food. In [14], an e-nose has been implemented to detect LPG leakage. Moreover, gas sensor array-based e-nose has been used to detect milk spoilage [15]. In [16], the potential application of an e-nose to monitor food spoilage has been shown. Also, an e-nose has been successfully implemented for several food items [17].

The development of gas sensors uses materials like NiO, SnO₂, Fe₂O₃ and ZnO [18], which are common metal oxide semiconductor gas-detecting materials. MOS sensors have long-lasting life and short sensing response time, operated over elevated operating temperature. MOS are inexpensive, solely sensitive to volatile gases and vapours, compact, and lightweight. Tin oxide, zinc oxide, titanium oxide, indium oxide, tungsten oxide, cerium oxide, copper oxide, and composite MOS for gases/odors detection are a few representative examples of the many metal oxide semiconductors. Metal Oxide gas sensors (MOX) are available at low cost, long lifetime, and fast response but relatively poor selectivity, drift in performance, and sensitivity to background gas. Schematic internal circuit diagram of low-cost SnO₂ gas sensor element has been shown in Figure 2. The electrical resistance of the gas sensing element as shown between A-B changes when it is exposed to certain concentration of the VOC, while the heater heats the sensing element. Output of the gas sensor is then taken against a load resistance, as a voltage divider bias circuit.

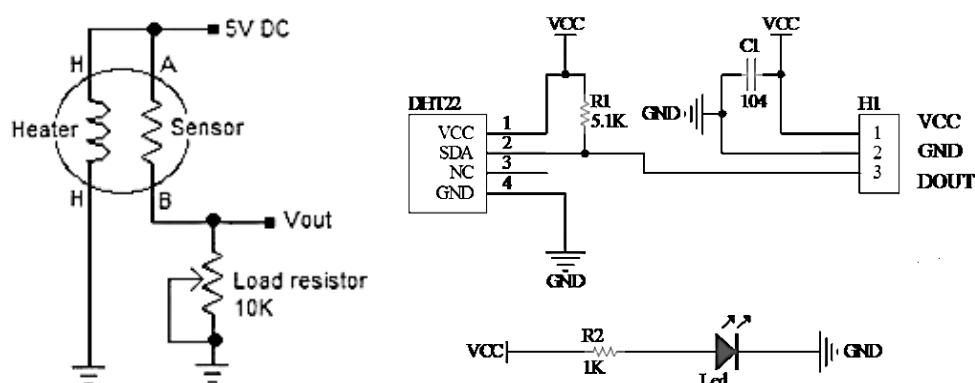


Figure 2. Schematic Internal Circuit Diagram of MQ Sensor and DHT 22

In this paper, we have developed a physical gas sensor system using a low-cost MQ series of the gas sensor array with six sensor elements, one temperature and one humidity sensor. The basic block schematic of the proposed e-nose is presented at Figure 3.

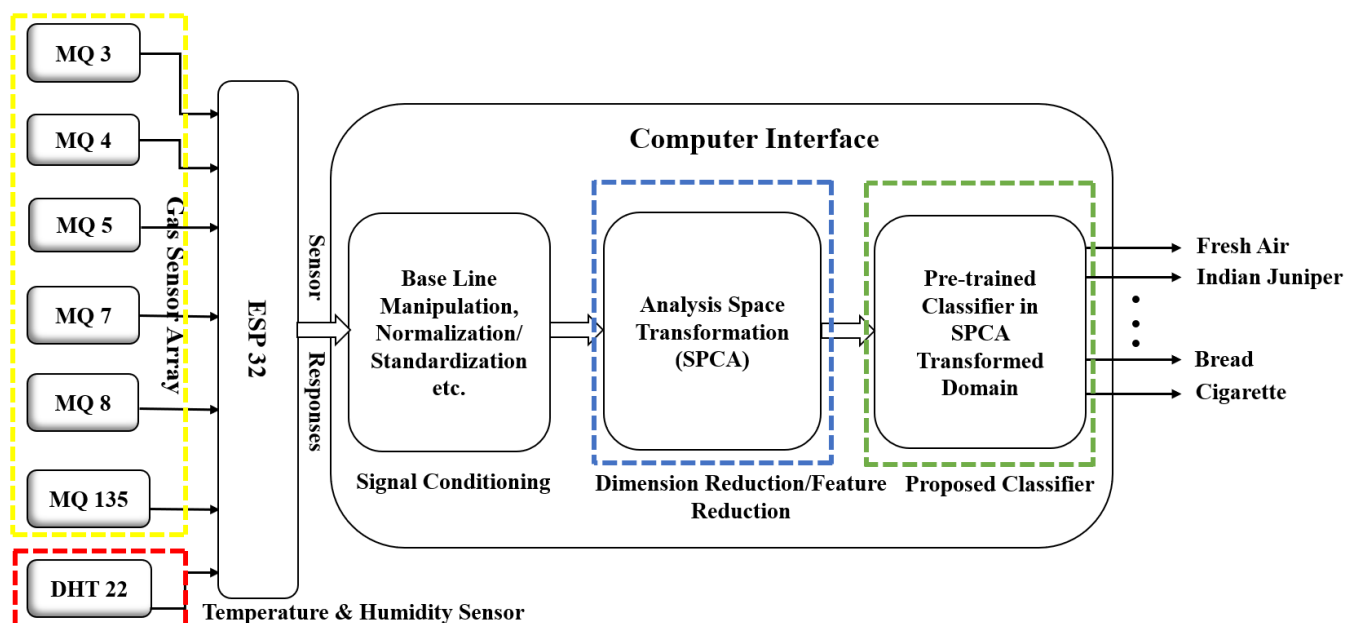


Figure 3. Block Schematic diagram of the proposed E-nose trained in the SPCA Transformed Analysis Space

Here, in this block schematic, we have added two more processing steps, i.e. the blocks shown in blue and green dashed lines, in Figure 3. The Blue block is a SPCA transformation block which transforms the raw sensor responses into SPCA-transformed analysis space. In this analysis space the data is clearer and creates very good clusters and a simpler ANN (shown in dashed green block) is trained using the SPCA transformed data. Rajput et al. (2010) had used a four-element gas sensor array for four hazardous gases and achieved all correct classification of the test samples [6].

In this proposed work, we have further optimized the gas sensor system to have six-element gas sensor array, hereafter called as High-performance Gas Sensor System (HP-GSS). We have trained and tested proposed gas sensor system over seventeen different types of VOCs, gases and smokes. The considered 17 household VOCs/gases/odors are presented in Table 1. Using a computer interface, we have captured the real-time gas sensor responses and segregated the same as training and testing dataset. Further, the testing dataset has not been used for training of the classifier.

Table 1. Distribution of Samples in training and testing dataset

S. No	Raw Materials	Training Set	Testing Set	Total Sample Size	Data Collection Time (Min.)	Class
1	Ambient Air	145	5	150	15	1
2	Indian Juniper	145	5	150	15	2
3	Mosquito Coil	145	5	150	15	3
4	All Out Liquid	145	5	150	15	4
5	Camphor	145	5	150	15	5
6	Zed-Black	145	5	150	15	6
7	Millennium	145	5	150	15	7
8	Heritage	145	5	150	15	8
9	First Spray	145	5	150	15	9
10	Harpic	145	5	150	15	10
11	Phenyl	145	5	150	15	11
12	Acid	145	5	150	15	12
13	LPG	145	5	150	15	13
14	Cow Milk	145	5	150	15	14
15	Bread (Wheat)	145	5	150	15	15
16	Pigeon Pea	145	5	150	15	16
17	Cigarette (Gold flake)	145	5	150	15	17
	Total	2465	85	2550	255	

Further, we have transformed the data in some of the popular transformation spaces viz. PCA, SPCA, LDA, KPCA and ICA. These pre-processing methods are employed to transform the raw sensor data into a domain where the data shows compact and well separated clusters. The proposed e-nose has been planned to be simple, portable, and affordable. Our used sensor elements are made of tin oxide with dopants like zinc oxide (ZnO) etc., and are generally cross-sensitive [5], [19]. In the proposed transformation domain and using a six-element sensor array and ANN-based pattern recognition techniques, we have identified all the 17 types of considered gases/odors (including the ambient air responses as one of the classes), accurately. Further details on PR procedures can be found in [5].

The physical gas sensor node fabricated for this experiment is depicted in Figures 5(a) - (c). Further, the following are the salient features of the proposed work:

- Physical environmental parameters: The proposed gas sensor node generates signature patterns of the gaseous ambience through an array of six non-selective gas sensing elements of tin-oxide, temperature and humidity sensor in real-time.
- Independent operation: The proposed system can be used for time-stamped real-time ambient data capturing with high accuracy and in any preferred time interval between two successive readings.
- Standalone: A DC-to-DC buck-converter with 12V, 18Ah battery provides increased mobility and portability. The overall weight of the gas sensor system node is significantly portable.
- Replaceability: All the sensors used in our prototype can be replaced independently without any difficulty, providing an advantage in real-field situations.
- Low-cost: This prototype has been developed using low-cost, high-performance ESP 32 Microcontroller [20], DC-to-DC buck-converter, MQ series gas sensors [21], and DHT-22 temperature and humidity sensor [22], commonly available online.
- Computer Interface: A simplistic programming helps capture the real-time ambient data in .csv format, and python 3.9 has been used for data pre-processing, PCA-based analysis space transformation and classifier design.

This paper has been organised into five sections. In section 1, the authors have provided related literature-based justification and motivation for the proposed research work. Section 2 detailed the sensor node prototype, data pre-processing and PCA transformation procedures. In section 3, the methodology towards high-performance classifier design has been given. Results

and discussion are then presented in section 4. The conclusions and future works are then drawn in section 5. Furthermore, all the abbreviations or acronyms are listed in Table 1 used throughout the articles.

2. Materials and Methods

In this paper, we have implemented a HP-GSS for use in smart homes for detection and monitoring of various VOCs, gases and odors, generated in the household sector during various daily chores such as worship, safety & security for hygiene, LPG leakage detection, cooking and smoking. As introduced in the previous section, the household sector has mainly been less attended to, whereas we spend our life's most precious hours at home itself.

Accordingly, in this work, we are motivated to develop an HP-GSS which can work over a wide range of gases/odors usually found in household ambience, especially in India. In this experiment, the most common 17 gases/odors have been considered, including the baseline composition of the ambient air as one of the classes. They are (i) released from various carcinogenic burning of food items (milk, wheat-bread, pigeon pea), (ii) smoke of different brands of incense sticks (Indian Juniper, Zed-Black, Millennium, Heritage, first-Spray), and (iii) miscellaneous aromas and gases emanated from house cleaning disinfectant (phenyl, Acid, Harpic), Camphor, mosquito repellent brands (All-out, Mosquito-coil), (Liquefied petroleum gas) LPG, Cigarette.

2.1 The High Performance – Gas Sensor System (HP-GSS) and Data Acquisition

We have developed a high-performance gas sensor node by interfacing six tin-oxide based gas sensor elements and one temperature and humidity sensor, in the form of an array, with a microcontroller. Further details of the aforesaid sensors and their connection pins have been given in Table 2.

Table 2. Component's List in Sensor Array with Description

S.No	Sensor Name	I/O Pin	Target Gas/Odor	Detection Ranges (PPM)
1	MQ 3	32	Alcohol, Ethanol, Smoke	25 – 500
2	MQ 4	33	Methane, CNG	300-10000
3	MQ 5	34	Natural Gas, LPG	300-10000
4	MQ 7	35	CO	10 - 500
5	MQ 8	36	Hydrogen	100-10000
6	MQ 135	39	Air quality	10 - 1000
7	DHT 22	25	Temperature and Humidity	-40 – 125 (°C)

The aforesaid microcontroller is a 30-pin ESP 32-Wroom, which operates at +5V Vcc while it's the sensor input at the GPIO pins are rated at 3.3V. Therefore, a voltage divider is used to translate the real-time analogue output of the gas sensing element (V_{GS}) operated in 5V range to the 3.3V input voltage range of the GPIO pins, as per the equation (1), where $R_1=2K\Omega$ and $R_2=3.9K\Omega$ has been taken.

$$V_{in(GPIO)} = \frac{V_{GS} \times R_2}{R_1 + R_2} \dots \quad (1)$$

Further, the main power supply has been obtained from a 12V, 18A Lead-acid Sealed Maintenance Free (SMF) battery for added portability. A DC-DC buck converter has been used to generate a 5V, 2.5A output from the SMF battery to power the microcontroller and the gas sensing elements. The gas sensing elements have been separately powered by the 5V, 2.5A supply, because each GPIO pin is rated at 3.3V, 40mA, and each gas sensing element operates at 5V, 150mA, requiring a total of 900mA current at 5V, for the considered six elements. The DHT22 temperature and humidity sensor has also been powered from this 5V, 2.5A supply, to retain the simplicity of the circuit. The ground terminals of each of the sensors, the microcontroller, and the DC-DC converter have been kept common with the ground terminal of SMF battery to ensure same voltage references. Electrical rating of all the components used in this circuit have been presented in Table 3.

Table 3. Electrical ratings of the components used in the circuit

Components	Input Voltage	Power Ratings
ESP 32	5V	130mA
DC-DC Buck converter	5V	2.5A
ESP32 GPIO pins	3.3V	40mA
DHT22	3-5V	2.5mA
MQ Sensors x 6	5V	150mA

The block schematic of the circuit as implemented on the PCB has been shown in figure 4. For programming the microcontroller, we have used Arduino Integrated Development Environment (IDE) by using various libraries viz. DHT.h, Adafruit_Sensor.h, analogRead(sensor name), dht.readHumidity(), dht.readTemperature(), Serial.begin(115200) interfaced with a computer running the IDE and Python 3.9. We have also used various Python libraries viz. Pandas, NumPy, ScikitLearn, Matplotlib and Seaborn. Among these libraries, Pandas is useful for data manipulation (data-frame creation), NumPy for arrays. Matplotlib and seaborn are useful for creating static, animated and interactive visualisation while SciKitLearnis useful for developing the classifier model.

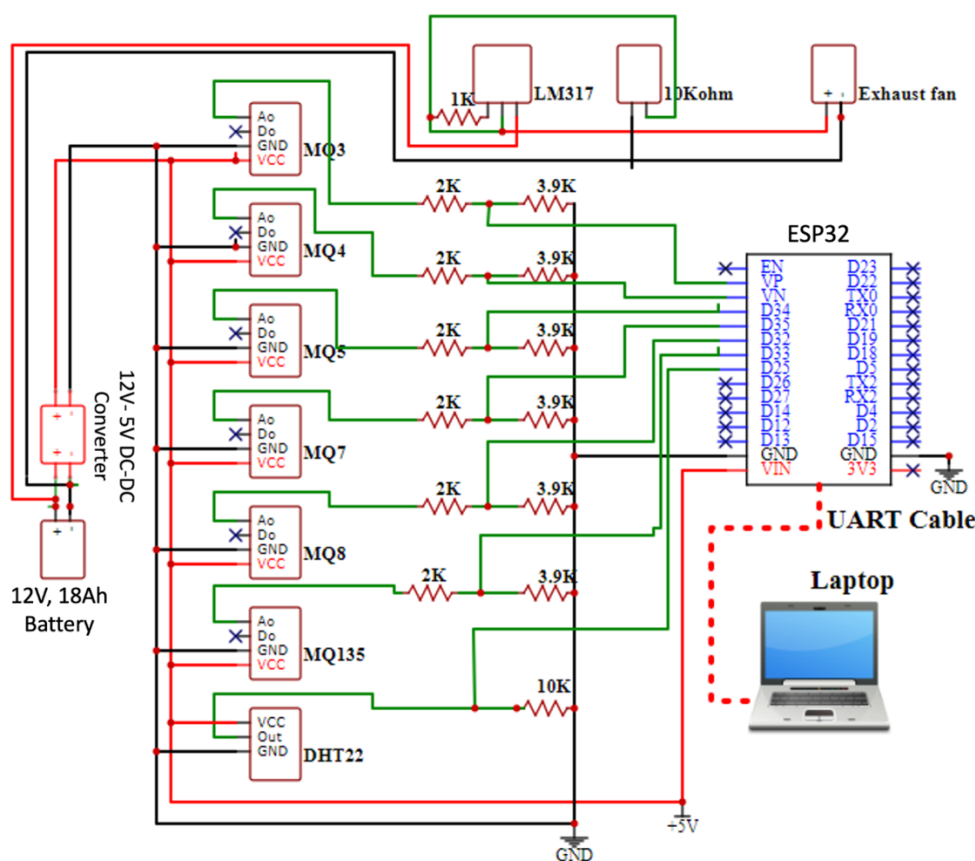


Figure 4. Schematic view of PCB design

The Sensors were placed on a PCB inside a gas chamber for operational purposes of gas exposure and sniffing. The chamber’s internal dimensions were 30 cm × 21 cm× 12 cm (L × W × H), with a total interior volume of 7.560 L (7650 cm³) and is made of 6 mm thick-glass and needful connections were made according to the circuit. Inlet and outlet holes have been created on two opposite side walls of the chamber. For proper flow of gases/odors, an exhaust fan of 120 mm diameter rated at 12V, 0.25A has been placed at the chamber’s outlet hole. This fan is directly powered by the primary 12V,18Ah SMF

battery. The speed of this exhaust fan is controlled by using a transistor LM137-based voltage regulator, as shown in Figure 5 (c). The physical prototype of the proposed gas sensor node is shown in Figure 5 (a) - (c).

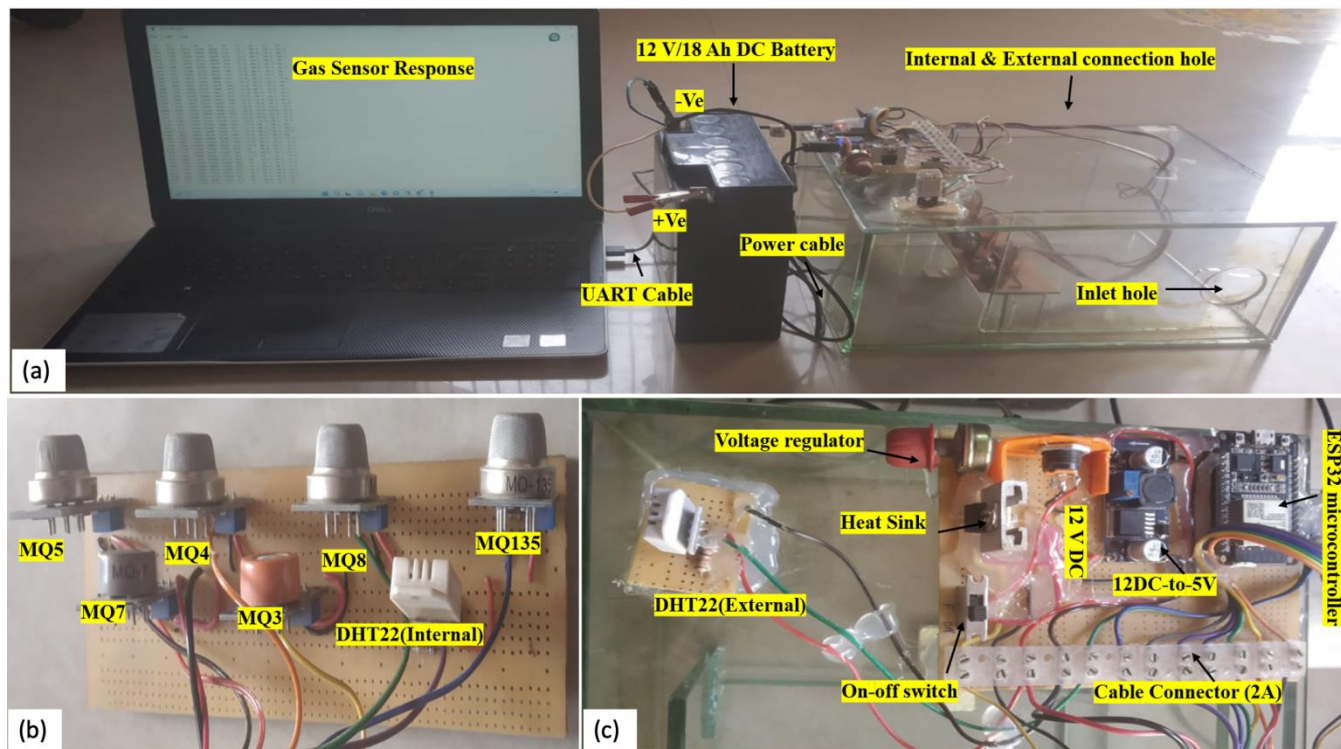


Figure 5. Prototype of the proposed High-Performance Gas Sensor System (a) Complete Setup (b) Sensor Array (c) Top-view

As shown in Figure 5(a), the proposed gas chamber is interfaced with a computer. Initially, the gas chamber is purged with the ambient air to obtain the baseline signature pattern under steady-state conditions of the sensor array responses for 30 minutes. Once the sensor array responses become steady, the array is exposed by one of the considered classes of gases/odors for 15 minutes at a rate of ten samples per minute in the comma-separated values (.csv) file format. After 15 minutes of exposure, the gas chamber is again purged with ambient air until the sensor responses return to the baseline signature patterns. This process is repeated for all the considered classes of VOCs, gases/odors and the raw sensor array response dataset is captured in its totality. Details of the captured dataset have been presented in Table 1.

2.2 The Concept and Approaches for Analysis Space Transformation

The raw sensor array dataset as captured during the HP-GSS exposure to the considered 17 classes of VOCs, Gases/Odors is analysed graphically, by observing its scatter plot, as shown in figure 8 (a). It has been observed in the published literature that when the clusters belonging to various classes are overlapping, a classifier trained using such sensor responses will perform poorly. In such cases, transforming the raw sensor responses in another suitable domain, called “analysis space” can be found and a classifier training in this domain can outperform [6], [23]-[27].

Accordingly, in this paper, we have explored various popularly effective analysis spaces domains such as Linear Discriminant Analysis (LDA) [8], Kernel Principal Component Analysis (KPCA) [11], Independent Component Analysis (ICA) [6], Principal Component Analysis (PCA) and Standardised Principal Component Analysis (SPCA) [6]. Theoretical details of each of the methods can be found in the given references. After comparison with the considered domain transformation methods, the SPCA based analysis space transformation domain has been found effective.

The SPCA based analysis space transformation method has been described in detail, as under:

The Process of transformation of Raw Gas Sensor Responses into SPCA domain

The raw sensor response dataset is first standardised for zero mean and unit variance. Then its covariance matrix is calculated to obtain its eigenvectors and eigenvalues. The highest eigenvalues and its associated eigenvector projects the first principal component (PC) along which the transformed dataset shows highest variance. Similarly, other principal components are also determined by taking next higher eigenvalues and eigenvectors. The information contained in each PC is obtained by normalising the eigenvalues and obtaining its percentage. The process of transforming the raw sensor array responses into SPCA analysis domain has been depicted in Figure 6.

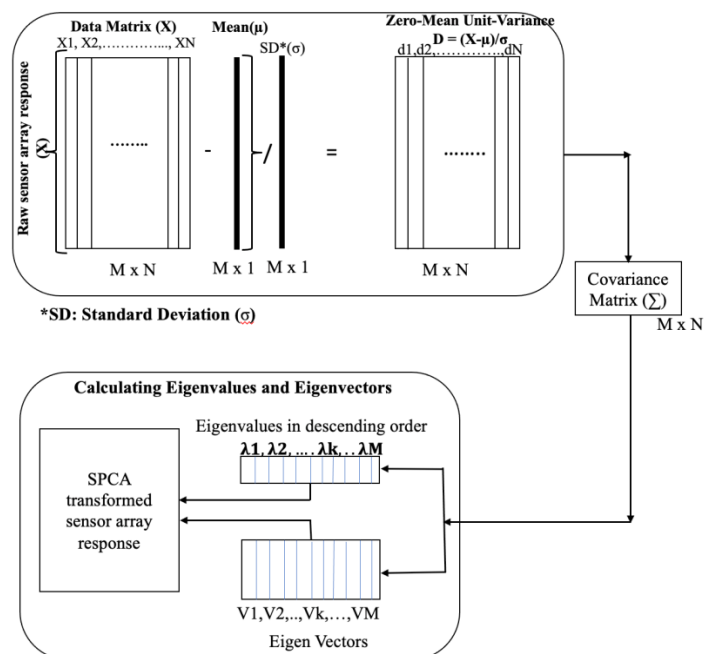


Figure 6. Graphical representation of the process of transformation of raw sensor responses into SPCA domain

2.3 Quantitative Details of the SPCA Transformed Sensor Array responses

The raw sensor array responses consisted of 2550 steady state responses, where 2465 samples were used for the training of the classifier. The remaining 85 samples were separated out for the testing purposes as described at Table 1. The 3D scatter plots for first 3 PCs of raw sensor responses and SPCA transformed sensor array responses have been shown in Figures 7 (a) - (b), for the sake of graphical visualisation. It can be observed in Figure 7 (a) that the clusters belonging to various classes are significantly overlapping (only a 50.53% variance in PC1). In Figure 7 (b), the 3D scatter plot the SPCA transformed sensor array responses have been shown. In the SPCA transformation domain, well separated clusters can be seen with a variance on PC1 was increased to 55.46%. It can therefore be expected that a classifier trained in the SPCA transformation domain should outperform.

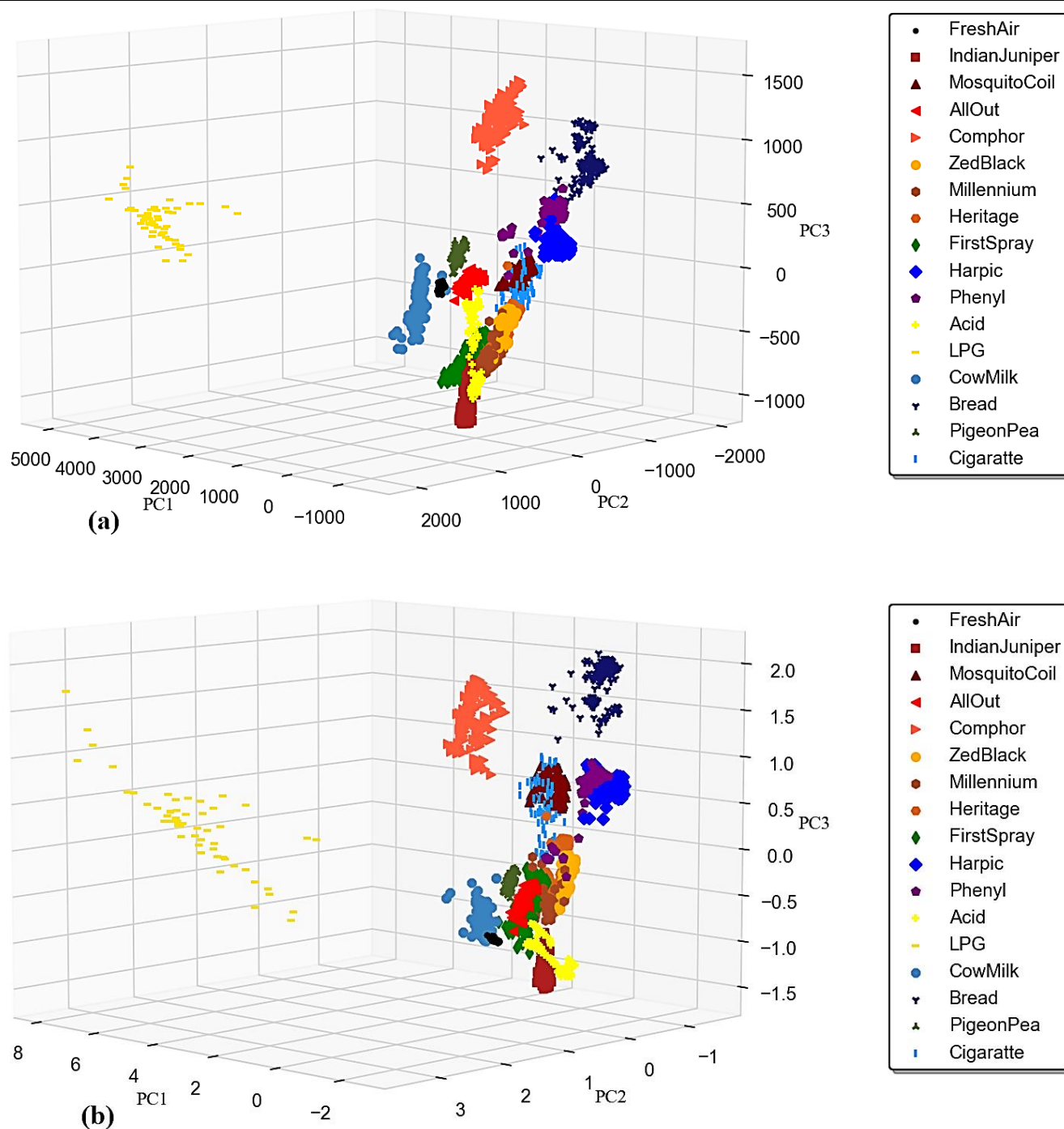


Figure 7. 3D scatter plots of first 3 PCs of sensor array responses in (a) raw and (b) SPCA transformed domain

2.4 Design of Classifiers using the SPCA Transformed Sensor Array responses

Once we have obtained the SPCA transformed details sensor array responses, we have trained three types of popular classifiers for classification of sensor responses in 17 considered classes. These classifiers are Support Vector Machines (SVM) [7], [8], Random Forest (RF) [8] and Multi-layer Perceptron (MLP) [6]. Support Vector Machines (SVMs) have been widely applied to pattern classification problems and nonlinear regressions. SVM is a fast and dependable classification algorithm that performs very well with a limited amount of data to analyse. Much of the computation is spent on tuning two important parameters, i.e., γ and C . γ is the parameter related to the span of an RBF kernel: the smaller the value is, the

wider the kernel spans. C controls the trade-off between the complexity of the SVM and the number of non-separable samples. A larger C usually leads to higher training accuracy. In our case the values of γ and C for the best performing SVM classifier were 0.9 and 1.0. Further, Random Forest (RF) classifier combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fuelled its adoption, as it handles both classification and regression problems. RF is a supervised machine learning algorithm that builds decision trees on different samples and takes their majority vote for classification and average in case of regression. The main limitation of random forest is that a large number of trees can make the algorithm too slow and ineffective for real-time predictions. RF is fast to train but quite slow to create predictions once they are trained. In our case, the best performing RF classifier has used 5 estimators (trees) with the gini criterion.

The most popular kinds MLPs are feed-forward networks consisting of neurons. The input, hidden, and output layers are adjacent in feed-forward ANN models. Multiple neurons are present in each layer. MLP between the inputs and outputs creates a mapping function. During the learning process, the input neurons pass the training input data to the neurons in hidden layer and forwards it to the subsequent hidden layers and output layer. Each neuron's weighted inputs and bias in the hidden layer are used to perform mathematical computations. The proper activation function of each neuron is then used to map the cumulative response, and an output is obtained. In our case, the best performing MLP classifier used single hidden layer with 11 neurons. The activation function used has been Relu and the learning rate was taken as 0.001. It was training in 1500 epochs using adam optimiser. The MLP classifier has outperformed the classification for all the 17 classes as considers for the considered 85 testing samples.

2.5 The Proposed Modular Architecture

Our proposed modular architecture has two modules viz. the SPCA transforming module and the MLP classifier module, as already explained and shown in Figures 8 (a) - (b). The SPCA transformation has been applied to standardised datasets in its first module. Such obtained principal components have been used as the input features for the subsequent module. In this second module, the MLP classifier has been training using the training data already transformed in the SPCA domain. Furthermore, the duly trained MLP classifier was tested using 85 unknown samples from the testing data. In this architecture, we have used a single hidden layer consisting of 11 neurons.

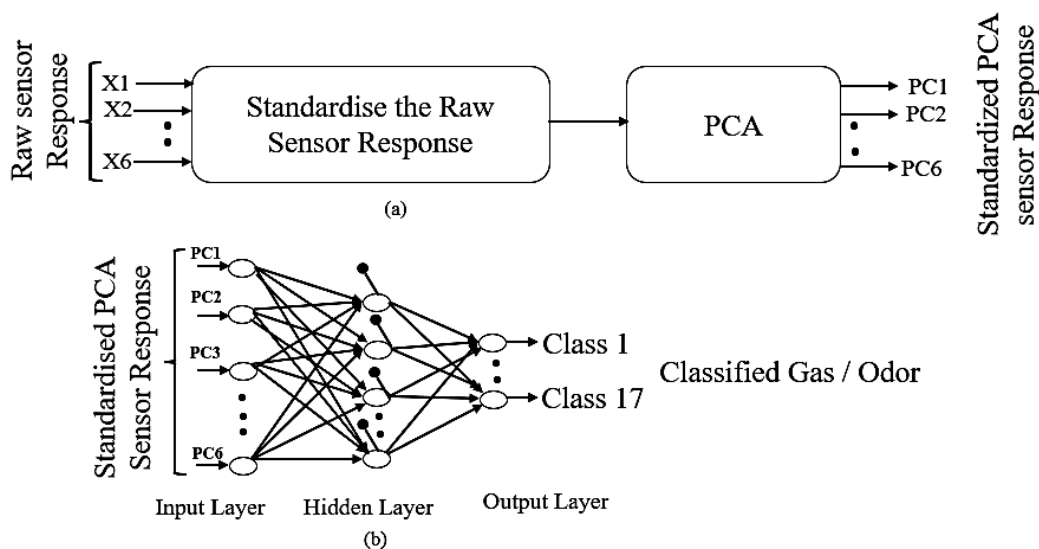


Figure 8. (a) Module I for SPCA transformation, (b) Module II for the MLP classifier.

The corresponding MLP classifier was trained using Levenberg—Marquardt (LM) training algorithm. All the classifiers were trained using a 5-fold cross-validation.

3. Results and Discussion

3.1 Laminar Air Flow using the variable speed of the exhaust

Laminar airflow allows air flow with a uniform velocity in straight lines parallel to each other. To minimize contamination, a laminar airflow system strives to keep the flow uniform and eliminate turbulence. While turbulent flows are unstable and

unpredictable, laminar flows are streamlined and uniform. The flow behaviour substantially alters if the flow is turbulent compared to laminar. The gas molecules do not get mixed up with each other while travelling. In this gas chamber, we have placed the sensor array PCB at three different positions viz. position # 1(5 cm), position # 2(15cm) and position # 3(25cm) from the right wall of the gas chamber, as shown in Figure 9. The device uses inwards air flow through sensor arrays (1, 2, and 3). We have observed that the sensor response of sensor array position-2 (middle position of the sensor array PCB) is steadier compared to sensor array position-1 (first sensor array) and sensor array position-3 (last sensor array).

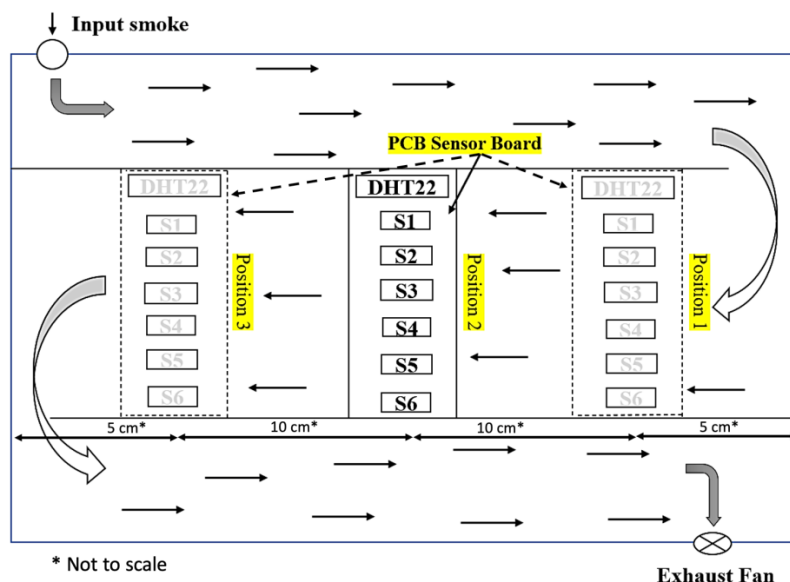


Figure 9. Pictorial representation of airflow in the gas chamber

3.2 Performance Analysis

The performance of the MLP classifier with raw sensor responses has been obtained and considered as a baseline for comparison. We have used SPCA transformation domain as a data pre-processing stage to enhance the variance in the direction of PCs.

The performance of the MLP classifier trained using the raw sensor array responses and trained using the sensor array responses in the SPCA transformation domain are shown in Figures 10 (a) - (b).

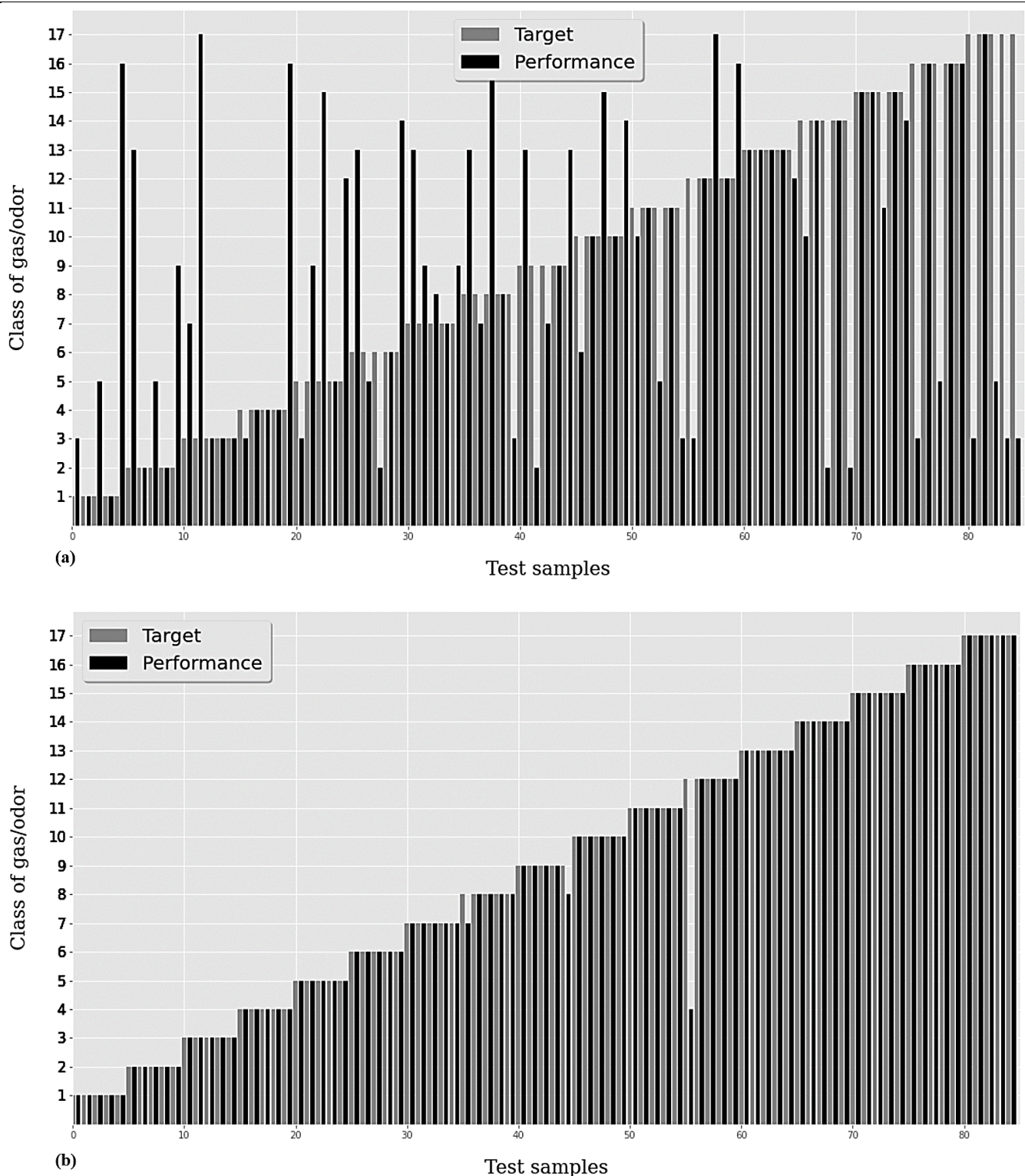


Figure 10. Performance of MLP classifier trained in (a) Raw and (b) SPCA transformation domain

Moreover, the efficacy of obtained results has also been shown using several confusion matrix-based parameters such as accuracy, precision, recall, F1-score, sensitivity, specificity, and G-mean determined using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) scores. The following equations are used to calculate these parameters.

$$Accuracy = \frac{TN + TP}{N} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (4)$$

$$F1_score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Where N = total number of samples

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

As shown in Figure 10 (a), the classification accuracy of 65.88% is achieved for the 85 unknown test samples. On the other hand, the MLP classifier trained in the SPCA Analysis space achieved a classification accuracy of 96.47%. The precision, recall and F1-score for the classification performance has been 96%, 95% and 95%, respectively.

3.3 Quantitative Analysis of the Results

While classifying the considered VOCs, gases/odors, we have compared the prediction error in each sample for the test samples. The error in correct classification was calculated using the popular error metric, i.e., Mean Squared Error (MSE). The error range ranges from 1.35×10^{-6} to 2.15×10^{-2} , providing an average MSE of 1.42×10^{-3} for 85 unknown test samples. The class wise mean squared error is shown in Figure 11.

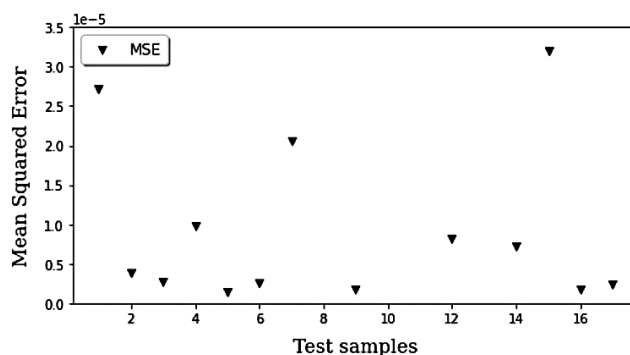


Figure 11. Error Plot on classification results of pre-processed dataset

For comprehensive visualisation of high-performance results, confusion matrix corresponding to the pre-processed dataset has been shown in Figure 12.

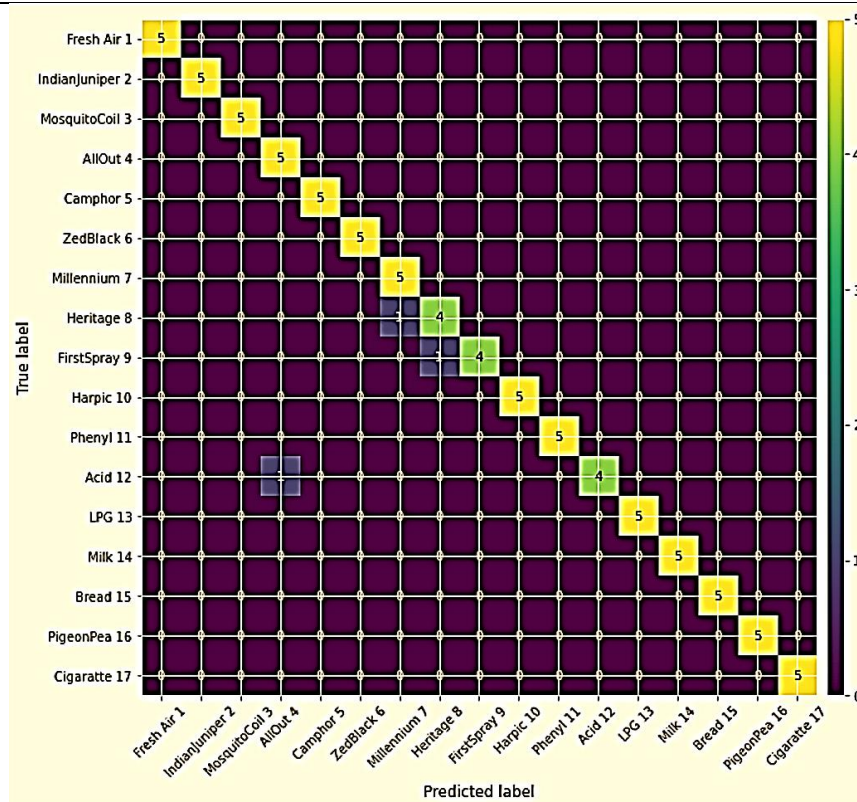


Figure 12. Confusion-Matrix for the MLP classifier training in SPCA transformation domain

A comparative performance of the three types of classifiers training in the considered five transformation domains have been presented in Table 4.

Table 4 Comparative performance of classifiers in various transformation domains

Classifier's Performance	Analysis Space Transformation Domains				
	ICA	KPCA	LDA	PCA	SPCA
MLP	87.45	90.11	95.56	65.88	96.47
SVM (rbf)	91.46	90.80	94.34	94.12	87.15
RF	95.25	94.99	95.52	90.59	93.29

4. Conclusions and future work

In this paper, we have developed High-Performance Gas Sensor System (HP-GSS) which can detect and monitor various VOCs, gases/odors as found in the domestic ambience emitted during daily chores (e.g., worship, safety & security for hygiene, LPG leakage detection, cooking and smoking). Our proposed HP-GSS incorporates a novel analysis space transformation approach implemented using MLP Classifiers. We have successfully demonstrated the classification of 17 different classes of VOCs, gases/odors. It can provide a quick, affordable, and efficient strategy in gas sensing-based domestic applications. Mimicking our proposed methodology, the developed HP-GSS are efficiently deployable in real-world scenarios. Being a modular approach, it can be extended over may other complex VOCs, gases and odors. We have demonstrated that simpler MLPs in transformed analysis spaces can outperform and can be designed in a generic manner. Our proposed HP-GSS is scalable and flexible to independently use any choice of pre-processing and pattern-recognition technique. Also, it provides crucial facilities in the area of intelligent systems, such as being portable, lightweight, and affordable. Additionally, the developed HP-GSS can be integrated with IoT enabled wearable sensor system. We have covered several contexts of air ambience in a smart home while demonstrating HP-GSS for detecting and monitoring the considered gases/odors.

ORCID iDs

Kanak Kumar  <https://orcid.org/0000-0002-1152-0680>

Navin Singh Rajput  <https://orcid.org/0000-0002-1650-011X>

Acknowledgements

This work was supported in part by the NCC Laboratory, Department of Electronics Engineering, IIT(BHU), India, under Grant IS/ST/EC-13-14/02 and I-DAPT HUB Foundation, IIT(BHU), India, under Grant R&D/SA/I-DAPT IIT(BHU)/ECE/21-22/02/290.

References

- [1] Jongyoon Park and Hitoshi Tabata, "Gas Sensor Array Using a Hybrid Structure Based on Zeolite and Oxide Semiconductors for Multiple Bio-Gas Detection", *ACS Omega* 2021, 6, 21284–21293, DOI:10.1021/acsomega.1c01435
- [2] Wenshen Jia, Gang Liang, Zhuojun Jiang, and Jihua Wang, "Advances in Electronic Nose Development for Application to Agricultural Products", *Food Analytical Methods*, 2019, 12, 2226–2240, DOI:10.1007/s12161-019-01552-1
- [3] Samaneh Matindoust, Majid Baghaei-Nejad and Mohammad Hadi Shahrokh Abadi, "Food quality and safety monitoring using gas sensor array in intelligent packaging", *Sensor Review*, 36/2, 2016, 169–183, DOI 10.1108/SR-07-2015-0115
- [4] K.I. Arshak, C. Cunniffe, E.G. Moore and L.M. Cavanagh, "Custom Electronic Nose with Potential Homeland Security Applications", *SAS 2006 IEEE Sensors Applications Symposium Houston, Texas USA, 7-9 February 2006*
- [5] J. Watson, "THE TIN OXIDE GAS SENSOR AND ITS APPLICATIONS," 1984.
- [6] N. S. Rajput, R. R. Das, V. N. Mishra, K. P. Singh, and R. Dwivedi, "A neural net implementation of SPCA pre-processor for gas/odor classification using the responses of thick film gas sensor array," *Sens Actuators B Chem*, vol. 148, no. 2, pp. 550–558, Jul. 2010, DOI: 10.1016/j.snb.2010.05.051.
- [7] L. Xu et al., "Hybrid Gas Sensor Array to Identify and Quantify Low-Concentration VOCs Mixtures Commonly Found in Chemical Industrial Parks," *IEEE Sens J*, vol. 22, no. 13, pp. 13434–13441, Jul. 2022, DOI: 10.1109/JSEN.2022.3176049.
- [8] E. Oleneva, T. Kuchmenko, E. Drozdova, A. Legin, and D. Kirsanov, "Identification of plastic toys contaminated with volatile organic compounds using QCM gas sensor array," *Talanta*, vol. 211, May 2020, DOI: 10.1016/j.talanta.2019.120701.
- [9] Rabeb Faleh, Mehdi Othman, Sami Gomri, Khalifa Aguir, Abdennaceur Kachouri, "A transient signal extraction method of WO₃ gas sensor array to identify pollutant gases", 1530-437X (c) 2015 IEEE Sensors Journal, DOI 10.1109/JSEN.2016.2521578
- [10] A. M. Andrew, A. Zakaria, S. M. Saad, and A. Y. Md Shakaff, "Multi-stage feature selection based intelligent classifier for classification of incipient stage fire in building," *Sensors (Switzerland)*, vol. 16, no. 1, Jan. 2016, DOI: 10.3390/s16010031.
- [11] Y. Xu, X. Zhao, Y. Chen, and W. Zhao, "Research on a mixed gas recognition and concentration detection algorithm based on a metal oxide semiconductor olfactory system sensor array," *Sensors (Switzerland)*, vol. 18, no. 10, Oct. 2018, DOI: 10.3390/s18103264.
- [12] L. Zhao, X. Li, J. Wang, P. Yao, and S. A. Akbar, "Detection of Formaldehyde in Mixed VOCs Gases Using Sensor Array with Neural Networks," *IEEE Sens J*, vol. 16, no. 15, pp. 6081–6086, Aug. 2016, DOI: 10.1109/JSEN.2016.2574460.
- [13] Phunvira Chongthanaphisit, Thara Seesaard and Teerakiat Kerdcharoen, "Monitoring of Microbial Canned Food Spoilage and Contamination Based on E-Nose for Smart Home", 978-1-4799-7961-5/15/ ©2015 IEEE
- [14] S. N. PATIL, S. R. GHATGE and A. M. PAWAR, "Implementation of Electronic Nose to Detect LPG Leakage", *i-manager's Journal on Embedded Systems*, Vol. 9, No. 2, January - June 2021
- [15] Natnaree Phukkaphan, Tanthip Eamsa-ard, Chalisa Chairanit and Teerakiat Kerdcharoen, "The Application of Gas Sensor Array based Electronic Nose for Milk Spoilage Detection", 7th International Conference on Engineering, Applied Sciences and Technology (ICEAST) | 978-1-6654-4122-3/20/ ©2021 IEEE | DOI: 10.1109/ICEAST52143.2021.9426263
- [16] Geoffrey C. Green, Adrian D.C. Chan, and Rafik A. Goubran, "Monitoring of Food Spoilage with Electronic Nose: Potential Applications for Smart Homes", 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare, IEEE, DOI: 10.1109/PCTHEALTH.2009.5291419

- [17] Chi Tai Dang, Andreas Seiderer, Elisabeth André, “Theodor: A Step Towards Smart Home Applications with Electronic Noses”, In 5th international Workshop on Sensor-based Activity Recognition and Interaction (iWOAR '18), September 20–21, 2018, Berlin, Germany. ACM, New York, NY, USA, Article 4, 7 pages. <https://doi.org/10.1145/3266157.3266215>
- [18] RashiBorgohain, Reema Das, Biplob Mandal, VisittapongYordsri, ChanchanaThanachayanont and Sunandan Baruah, “ZnO/ZnS core-shell nanostructures for low concentration NO₂ sensing at room temperature”, IEEE Sensors Journal, 1558-1748 (c) 2018 IEEE, DOI: 10.1109/JSEN.2018.2851196
- [19] A. Chaturvedi, V. N. Mishra, R. Dwivedi, and S. K. Srivastava, “Response of oxygen plasma-treated thick film tin oxide sensor array for LPG, CCl₄, CO and C₃H₇OH.”
- [20] ESP-32 microcontroller datasheet, https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf
- [21] MQ sensors datasheet, <https://www.pololu.com/file/0J309/MQ2.pdf>
- [22] DHT 22 temperature and humidity sensor datasheet, <https://www.sparkfun.com/datasheets/Sensors/Temperature/DHT22.pdf>
- [23] Ashutosh Mishra, N. S. Rajput, and Guangjie Han, “NDSRT: An Efficient Virtual Multi-Sensor Response Transformation for Classification of Gases/Odors”, IEEE SENSORS JOURNAL, VOL. 17, NO. 11, JUNE 1, 2017, DOI: 10.1109/JSEN.2017.2690536
- [24] Shiv Nath Chaudhri, Navin Singh Rajput, Ashutosh Mishra, “A novel principal component-based virtual sensor approach for efficient classification of gases/odors”, Journal of ELECTRICAL ENGINEERING, VOL 73, 2022, NO2, 108–115, DOI: 10.2478/jee-2022-0014
- [25] Shiv Nath Chaudhri, Navin Singh Rajput, Saeed HamoodAlsamhi, Alexey V. Shvetsov and Faris A. Almalki, “Zero-Padding and Spatial Augmentation-Based Gas Sensor Node Optimization Approach in Resource-Constrained 6G-IoT Paradigm”, MDPI, Sensors 2022,22, 3039, DOI: 10.3390/s22083039
- [26] S. Srivastava, S. N. Chaudhri, N. S. Rajput, S. H. Alsamhi and A. V. Shvetsov, "Spatial Upscaling-Based Algorithm for Detection and Estimation of Hazardous Gases," in IEEE Access, doi: 10.1109/ACCESS.2023.3245041.
- [27] S. N. Chaudhri and N. S. Rajput, "Multidimensional Multiconvolution-Based Feature Extraction Approach for Drift Tolerant Robust Classifier for Gases/Odors," in IEEE Sensors Letters, vol. 6, no. 4, pp. 1-4, April 2022, Art no. 7001004, doi: 10.1109/LSENS.2022.3153832.