

Northumbria Research Link

Citation: Akbar, Muhammad Ali, Ait Si Ali, Amine, Amira, Abbas, Bensaali, Faycal, Benammar, Mohieddine, Hassan, Muhammad and Bermak, Amine (2016) An Empirical Study for PCA- and LDA-Based Feature Reduction for Gas Identification. IEEE Sensors Journal, 16 (14). pp. 5734-5746. ISSN 1530-437X

Published by: IEEE

URL: <https://doi.org/10.1109/JSEN.2016.2565721>
<<https://doi.org/10.1109/JSEN.2016.2565721>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/31467/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

An Empirical Study for PCA and LDA Based Feature Reduction for Gas Identification

Muhammad Ali Akbar, Amine Ait Si Ali, Abbes Amira, *Senior Member, IEEE* Faycal Bensaali, *Senior Member, IEEE* Mohieddine Benammar, *Senior Member, IEEE*, Muhammad Hassan and Amine Bermak *Fellow, IEEE*

Abstract— Increasing the number of sensors in a gas identification system generally improves its performance as this will add extra features for analysis. However, this affects the computational complexity, especially if the identification algorithm is to be implemented on a hardware platform. Therefore feature reduction is required to extract the most important information from the sensors for processing. In this paper, linear discriminant analysis (LDA) and principal component analysis (PCA) based feature reduction algorithms have been analyzed using data obtained from two different types of gas sensors i.e. seven commercial Figaro sensors and in-house fabricated 4x4 tin-oxide gas array sensor . A decision tree (DT) based classifier is used to examine the performance of both PCA and LDA approaches. The software implementation is carried out in MATLAB and the hardware implementation is performed using the Zynq system on chip (SoC) platform. It has been found that with the 4x4 array sensor, two discriminant function (DF) of LDA provides 3.3% better classification than five PCA components, while for the seven Figaro sensors two principal components (PC) and one DF show the same performances. The hardware implementation results on the programmable logic of the Zynq SoC shows that LDA outperforms PCA by using 50% less resources as well as by being 11% faster with a maximum running frequency of 122 MHz.

Index Terms— Feature reduction, Gas identification, PCA, LDA, Electronic nose, Zynq SoC

I. INTRODUCTION

Gas sensing platforms are widely used in various applications ranging from monitoring pollution to industrial applications. In 2006, the first electronic nose (EN) system was presented which provided different radial graphs for hydrogen (H₂) and carbon-mono-oxide (CO) but did not

provide any details for further classification [1]. Therefore, Victor et al. proposed in [2] an EN with tin-oxide based microarray, which can discriminate between several gases in air. However, the current commercialized EN systems suffer from the high cost of fabrication along with large size. Different research studies have been carried out to reduce the cost and size of EN systems. A micro-electro-mechanical system (MEMS) based EN system has been reported in [3] which reduces the power consumption to 23mW with the expense of sophisticated MEMS approach.

Aging and long term exposure to reactive gases can result in a change of the gas sensor properties [4]. The two most critical challenges for gas identification are the drift and non-selectivity of the sensors [5]. The problem of non-selectivity can be overcome by utilizing more than one sensor at a time such that each sensor shows different sensitivity or response to the target gas. Thus, a multiple-sensor approach is adopted in gas application to obtain different signatures of the gas at a time. For example, Guo et al. in [3], proposed a 4x4 array gas sensor in which each sensor provides a different response for any particular gas. Another approach to deal with the non-selectivity is the temperature modulation and fluctuation enhanced sensing (FES) which can also be used to identify the gas with only a single sensor [4]. In case of temperature modulation the sensor operates at different operating temperatures such that it provides different responses at a time. Similarly, in FES the noise spectrum of sensor is used to determine the chemical information of the target gas [6]. Liobet et al. utilized the concept of temperature modulation in [7] and proposed a single SnO₂ based gas identification system using discrete wavelet transform (DWT). Whereas Kish et al. in [8], used the power spectrum of the noise to determine the gases. However, both approaches of multiple-sensors and single sensor-modulation increase the dimensionality of the feature vector, thereby increasing the computational complexity [6][9]. A problem besides the computational complexity/cost with high-dimensional feature vectors is the “curse of dimensionality” [10]. The problem becomes more severe if the gas identification system is implemented on any hardware platform because of resource utilization and power consumption, both of which increase with the computational complexity. In addition gas recognition process requires a complicated training phase and frequent calibration, which is even harder to implement on a

This paper was made possible by National Priorities Research Program (NPRP) grant No. 5 – 080 – 2 – 028 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

M. A. Akbar, A. Ait Si Ali, A. Amira, F. Bensaali and M. Benammar are with Qatar University, Doha, Qatar (e-mail: ali.akbar@qu.edu.qa).

M. Hassan and A. Bermak are with School of Engineering, Hong Kong University of Science and Technology, Hong Kong.

dedicated hardware. Therefore, the aim of this research is to identify the best feature reduction approach between the PCA and LDA, which is suitable for hardware and software implementation and can be applied in any multi-sensing gas identification platform.

Feature reduction algorithms are used to reduce the data size while keeping enough information to be able to discriminate efficiently between the classes [11]. Different research approaches have already been presented for feature reduction like independent component analysis (ICA) [12], multidimensional scaling [13], etc. The two most common techniques for dimensionality reduction are principal component analysis (PCA) and linear discriminant analysis (LDA).

PCA is used to determine a subspace that increases the variance between data and thus dimensionality reduction can be achieved by projecting data in the new subspace [14]. Sophian et al. in [15], applied PCA in extracting features from pulse eddy current (PEC) responses. Moreover, a PCA based multi-lead analysis approach is used by Monasterio et al. in [16] to improve the estimation and detection of a cardiac phenomenon known as T-wave alternans (TWA) in biomedical applications. In EN systems, PCA is used along with the DT by Qingzheng and Bermak in [17]. In contrary, LDA uses an assumption that the data classes have an equal covariance structure with Gaussian distribution [18]. It reduces the distances, present with in-class and increases the distances between classes simultaneously. Chakrabarti et al. in [19] applied LDA for text classification. LDA is used for the classification of tumor data obtained from microarray by Dudoit et al. in [20]. In face recognition LDA is used in [21]. LDA is also used in gas identification by Ankara et al. in [22].

Most of the EN systems have been implemented either using a uniprocessor-based software approach or a hardware-based implementation approach like field programmable gate array (FPGA) which accelerates the slow software-based approach to meet the real-time requirements. However, with the emergence of novel platforms such as the Xilinx Zynq [23] which holds on a single chip, a processor and an equivalent of FPGA which allows an efficient and quick hybrid based implementation approach especially when associated with high level synthesis (HLS) tool. In a hardware software co-design implementation, computationally intensive blocks of the EN system can be executed on hardware, while the remaining non-complex tasks can be performed on a processor in a software manner. This approach will not only reduce the power consumption of the hardware, but also provides more space in the FPGA for other tasks related to hardware acceleration. The presented work is part of an ongoing project in which a low-power multi-sensing gas identification platform is being developed for gas identification using an array of tin-oxide gas sensors.

This paper compares the impact of both PCA and LDA based feature reduction approaches on a decision tree (DT) based classifier for gas identification systems. The gas data is extracted using two different types of sensors: an in-house fabricated 4x4 array tin-oxide gas sensor [3] and a set of 7

Figaro commercial sensors [24]. Moreover, two different properties of data extracted from the sensors are used for classification: the first, referred to as steady state (SS), which considers the absolute reading from the sensor, and the second, referred to as delta (Δ), which considers the recorded changes in sensor reading between exposure to air and target gas. A 5-fold cross-validation (CV) approach has also been adopted to verify the obtained results under different samples for training and testing operations. This will not only help to minimize the problem of overfitting but also supports to assess the statistical significance of the obtained results by using some statistical metrics like standard deviation (SD) and coefficient of variation (CoV). The software implementation is carried out using MATLAB and for hardware implementation the Zynq reconfigurable SoC platform has been used where the flexibility offered by these platforms has been exploited to re-program the classification algorithms.

The obtained result shows that with the SS of the data extracted from the 4x4 array sensor, DT successfully classifies 94.99% of the gases with first two discriminant function (DF) of LDA, which in case of PCA reduced to 91.66% using five principal components (PC). In terms of hardware resources and computational time, LDA uses on average 50% less resources than the PCA and is 11% faster with a maximum running frequency of 122 MHz.

The remaining sections of this paper are organized as follows. Section 2 covers the experimental setup. The PCA and LDA-based dimensionally reduction algorithms are described in Section 3. Section 4 is concerned with the achieved software and hardware results and their discussion. Section 5 concludes the paper.

II. EXPERIMENTAL SETUP

The gas data is extracted using two different types of sensors so that comparative analysis between PCA and LDA for gas identification can be performed under diverse environments. Moreover, different gases along with one common gas are used by both sensors for data extraction. This will verify the experimental results under both possible condition of similarity and dissimilarity used for data extraction. The use of different gases and sensor types provide diversity to our design approach, thereby validating the analysis results of PCA versus LDA. Hence, the conclusions obtained from this research provides the most suitable feature reduction approach out of PCA and LDA for gas identification, regardless of the type of sensor and nature of the gas. The overall system design for gas identification is shown in Fig. 1.

The data from gas sensor is extracted under a controlled laboratory environment, as shown in Fig. 1. The experimental setup comprises of three cylinders containing different premixed gases with specified concentrations and a separate air cylinder which is used for diluting the premixed gases and for flushing the sensors with pure air. Each experiment requires 10 different concentrations of each gas. Therefore, the mass flow controllers (MFC) are used to adjust the flow of gas and air according to the required concentration.

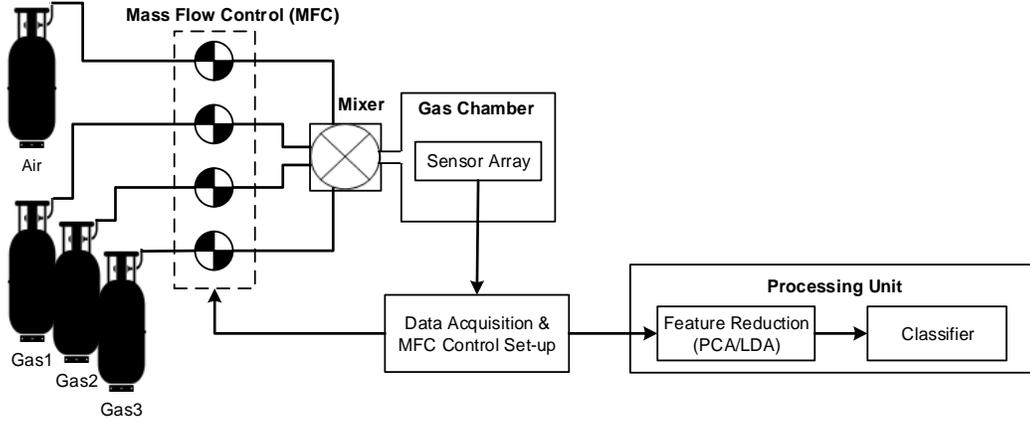


Fig. 1. Proposed investigation approach for gas identification system.

In order to extract the data, gas sensor array is placed in a glass chamber which has an inlet at one side and an outlet on the opposite side. The sensors are of resistive type; the change of sensor resistance with the target gas is transformed into a voltage change using a voltage divider network.

Thereafter, the voltage is read by a 12-bit analog to digital converter (ADC) of the data acquisition board which is interfaced to the computer. A computer software is used to automatically control the MFC for changing the gas concentration according to a pre-set schedule and log the sensors' responses at a suitable sampling rate. Each sensor type provides data for ten concentrations (20, 40, 60, 80, 100, 120, 140, 160, 180 and 200ppm in air) of different gases after which the process is repeated again for the same gas with the similar range of concentrations as before to obtain the test patterns which verify the classification algorithm. Hence, the obtained data are divided into two parts in which the first half of the data are used for the learning purpose and is named as training data, while the second half is used for testing and verification of the algorithm.

Consider $Gas_{i \times j}$; where i represents the type of gas (i.e. 1, 2, 3, ..., x) and j is the concentrations (i.e. 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200ppm in air). Let t_1 and t_2 be the time of air and gas exposure respectively. Both sets of sensors use identical approach for extracting the data which are illustrated in the flowchart shown in Fig. 2. It can be observed that before exposing the sensors to any new concentration of the gas, they need to be flushed with air to remove all traces of the previous experiment and bring the sensors back to the initial state. This is because the sensors are characterized by different response times, and in order to unify exposure time, test runs have been conducted to select suitable times t_1 and t_2 such that SS are ensured for the sensors. This approach helps in making a uniform experimental setup for a particular sensor.

The data extracted from the sensors are a combination of transient and SS, however, in most cases the SS or static values are highly recommended to make the variable time-invariant [25]. Therefore, the time intervals t_1 and t_2 are chosen long enough to ensure that the sensors' outputs reach

SS for air and gas, respectively. However, some sensors did not reach strictly stable state. Instead of waiting for a long period, the time of gas delivery and air injection phases has been set according to the needs. The sensor readings are extracted through the average of the last 5 samples of each sensor during each phase. This approach is referred to as a SS approach for the experiment.

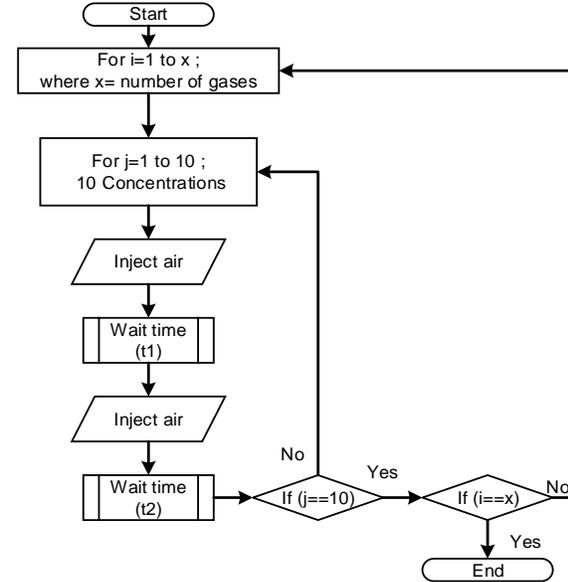


Fig. 2. Flowchart for data acquisition

Another approach to dealing with the sensor data is to use relative values with respect to air. In [25] a fractional different approach is defined, as shown in Eq. 1; where x is the sensor number and j is the gas concentration. However the division operation can result in increase in area overhead for hardware, therefore in this research the difference operator Δ which is defined in Eq. 2 is used.

$$F.D = \frac{(S_{xj}^{Gas} - S_x^{air})}{S_i^{air}} \quad (1)$$

$$\Delta = (S_{xj}^{Gas} - S_x^{air}) \quad (2)$$

A. 4x4 Array of tin-oxide based Gas Sensor

The 4x4 array gas sensor is used to extract data for 10 concentrations of three different gases, namely carbon-monoxide (CO), ethanol (C₂H₆O) and hydrogen (H₂). The layout and composition of the 4x4 array sensor are shown in Fig. 3(a) and (b), respectively. The air is passed through sensor for **t₁ = 750 s (seconds)** and then the sensor is exposed to the new concentration of gas for **t₂ = 250 s**. Therefore, the test at each concentration takes 1000 s and the overall time for 10 concentrations become 10,000 s.

The response of the sensor for 10 concentrations of ethanol gas is shown in Fig. 4(a). It has been noticed that the peak values of the graph represent the response of sensor to pure air and the base values or minima of the graphs are the respective response of sensor to each concentration of gas i.e. SS.

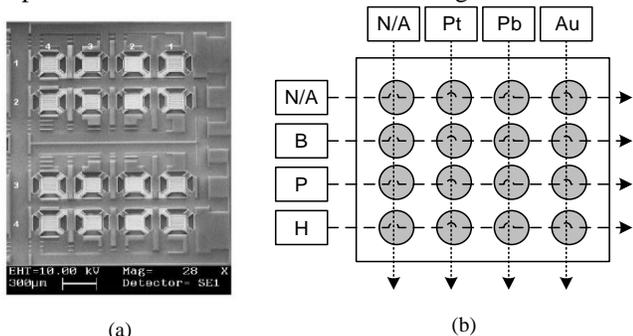


Fig. 3. 4x4 array sensor (a) layout [3] (b) composition of the sensor

B. Figaro Commercial Sensor

A set of seven different commercially available Figaro sensors, listed in Table I, are used to extract the data from four different gases, namely carbon-di-oxide (CO₂), hydrogen (H₂), ammonia (NH₃) and propane (C₃H₈). The exposure time for air remains the same for Figaro sensors (i.e. **t₁ = 750 s**), however, the time for gas exposure (**t₂**) increases to 500 s.

Although, for some gases, the sensors achieved SS before **500 s** like propane for which the SS is achieved at **250 s**. However, the maximum response time required by the sensors to attain SS for CO₂ is **500 s**, therefore all gases have been examined for **t₂ = 500 s**. The overall experimental time required for a single concentration of gas is equal to 1250 s and for 10 concentrations of each gas the value reaches 12,500 s. The response of the seven Figaro sensors for propane gas is shown in Fig. 4(b).

TABLE I
LIST OF COMMERCIAL SENSORS USED

| Sensor # | Sensor Model | Target Gas |
|-----------|--------------|---------------------------------|
| Sensor 01 | TGS 826 | Ammonia |
| Sensor 02 | TGS 2442 | Carbon-Mono-Oxide |
| Sensor 03 | TGS 2600 | Air contaminants |
| Sensor 04 | TGS 2602 | Volatile organic compound (VOC) |
| Sensor 05 | TGS 2610 | Liquefied petroleum (LP) gas |
| Sensor 06 | TGS 2611 | Methane |
| Sensor 07 | TGS 2620 | Solvent Vapors |

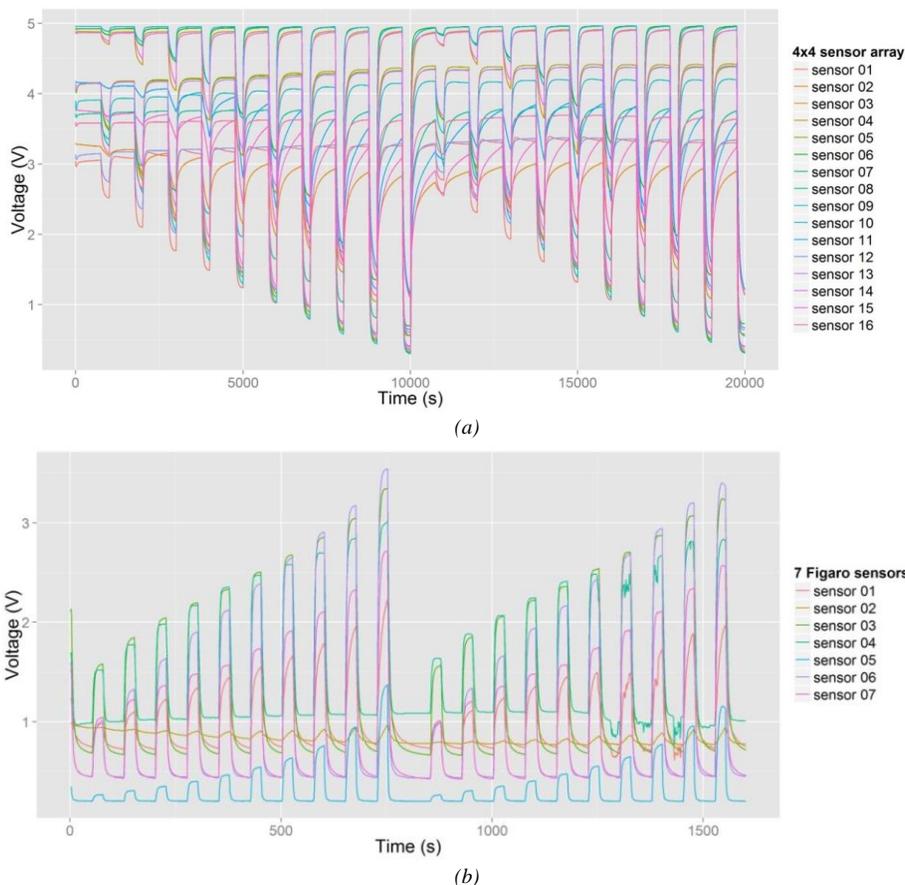


Fig. 4. Sample of sensor responses for (a) Ethanol by 4x4 Array sensor and (b) Propane gas by Figaro sensor; for 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200ppm in air

III. FEATURE REDUCTION ALGORITHMS

The data set extracted by each of the sensors is divided into two subsets named as training data (**D1**) and testing data (**D2**). Thus, both **D1** and **D2** are in matrix form with order of $m \times n$, where m is the number of samples and n is the number of sensors used. In this research m is a variable which depends on the sample size used for training and testing data, however n is equal to 16 and 7 for 4x4 array and Figaro sensors, respectively. The data subset **D1** is used during learning phase of the classifier, while **D2** is used for the validation and verification of PCA/LDA. In total 60 samples using the 4x4 sensor array and 200 samples from the seven Figaro sensors have been collected. The reason of having small data size is due to the fact that the data acquisition process is time consuming because it takes into consideration the gas and air injection into the chamber as well as reaching the SS and the baseline. Therefore the problem of overfitting in the training data can be avoided by properly cross-validating the classification models. One of the widely used technique for CV is the K-fold CV; it consists in partitioning the data into k equally sized subsets. The training and testing are performed k times where each time one of the k subsets is used for training and the other subset is used for testing. The average successful or error classification rate is computed by taking the mean of all the results obtained at each fold. In this work a K-fold CV approach has been adopted to check and verify the results for both PCA and LDA on a different set of training and testing data. K-fold CV is performed on the collected data with $k = 5$ which means the training and testing will be performed 5 times with 80% and 20% of the data, respectively. The results obtained from each fold are used to compute the average mean value for each reduction component and the corresponding SD using Eq. 3 and 4, respectively. The concluded results for each component “ c ” of PCA/LDA are statistically examined using the CoV which is obtained by dividing the corresponding SD by the average mean value as shown in Eq. 5. The obtained results are examined and discussed in Section IV.

$$\bar{x}_c = \frac{1}{K} \sum_{i=1}^K x_{i_c} \dots \dots \dots (3)$$

$$SD_c = \sqrt{\frac{1}{K} \sum_{i=1}^K (x_{i_c} - \bar{x}_c)^2 \dots \dots \dots (4)}$$

$$CoV_c = \frac{SD_c}{\bar{x}_c} \dots \dots \dots (5)$$

where x_i = Accuracy of each Fold;
 \bar{x} = Accuracy mean of all Folds;
 c = Corresponding component number
 and K = number of Folds

The two parameters which are applied to examine the data for gas classification are the SS and Δ . The comparative analysis between LDA and PCA is further deepened by taking one parameter at a time. In the beginning PCA and LDA based

feature reduction approaches are applied to the SS which are obtained when the sensor is exposed to the gas. The second performance matrix is analyzed on Δ value which is the absolute difference between the values of sensors obtained at air and gas, respectively.

A. Principal Component Analysis (PCA)

PCA is a statistical and unsupervised approach used for feature extraction and data compression [26] [27]. The purpose of PCA is to project the feature from high-dimensional to a new low-dimensional space where the derived axes known as principal component are having decreasing order of importance. The goal of PCA is to maximize the variance between data without considering class separation [28]. The steps for performing PCA in training and testing are summarized in the pseudo code shown in Algorithm 1. The mean computed from the training data will be used in the normalization during both training and testing processes as shown in Fig. 5 (a).

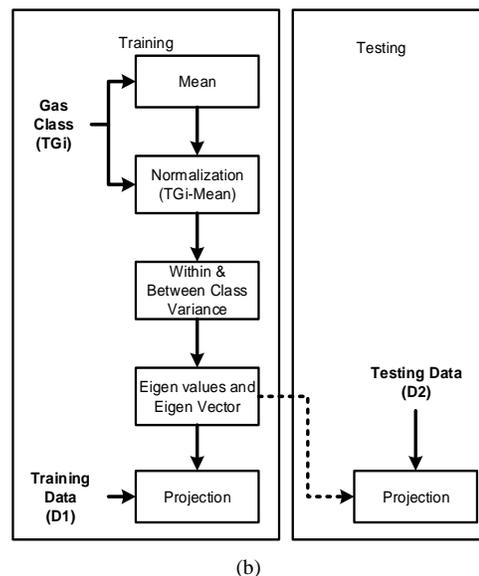
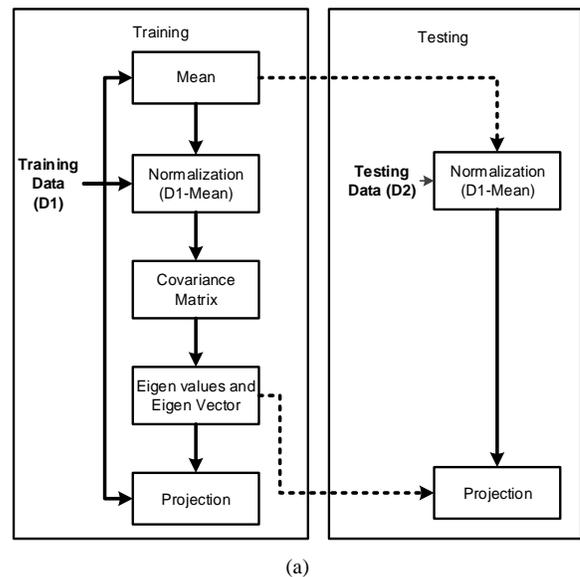


Fig. 5. Feature reduction algorithms (a) PCA (b) LDA

Algorithm 1: PCA Training and Testing**Required:** $D1_{m \times n}, D2_{m \times n}$ where $m =$ number of samples used $n =$ number of sensors**Training:**

- 1) $\vec{\mu} = [\mu_1 \mu_2 \dots \mu_n] \rightarrow$ mean of each sensor value in $D1$
- 2) $L_{m \times n} = (\sum_{i=1}^m \sum_{j=1}^n (D1_{i,j} - \mu_j)) \rightarrow$ Perform Normalization
- 3) $C_{m \times n} = \text{Cov}(L_{m \times n}) \rightarrow$ Covariance Matrix
- 4) [Ev, Eval] \rightarrow [Eigen Vector, Eigen value]
- 5) [Ev, Eval] = Eig($C_{m \times n}$)
- 6) $PCA_{\text{training}} = \text{Ev} \times D1_{m \times n} \rightarrow$ Projection

Testing:

- 1) $T_{m \times n} = (\sum_{i=1}^m \sum_{j=1}^n (D2_{i,j} - \mu_j)) \rightarrow$ Perform Normalization
- 2) $PCA_{\text{testing}} = \text{Ev} \times D2_{m \times n} \rightarrow$ Projection

B. Linear Discriminant Analysis (LDA)

LDA is the most commonly used supervised approach for feature reduction [29]. The basic function of LDA is to reduce the distances present within-class and increase the distances between classes simultaneously. In order to analyze the best reduction approach for the designed experimental setup, LDA is performed on the same data used for PCA. The overall technique is illustrated in the pseudo code shown in Algorithm 2. In contrary to PCA, the learning phase of LDA is not performed on $D1$ because LDA deals with class boundaries which cannot be identified in $D1$. Therefore, $D1$ is divided into classes such that data for each gas “ G_s ” is considered as a single class. The training-data obtained at different concentrations of G_s is stored as a T_{G_s} , where sub-script s is representing the type of gas. However, $D2$ is used directly for testing purposes as shown in Fig. 5(b).

Algorithm 2: LDA Training and Testing**Required:** $D1_{m \times n}, D2_{m \times n}, T_{G1}, T_{G2}, T_{G3}, \dots, T_{Gi}$ **Training:**

- 1) **Let $j_i \rightarrow$ number of samples of i th gas**
- 2) $\vec{\mu}_{T_x} = [\mu(T_{G1}) \mu(T_{G2}) \dots \mu(T_{Gx})]$ where $x = 1, 2, \dots, i$
- 3) $\mu = \sum_{x=1}^i \mu_{T_x} / i$
- 4) $\vec{C}_{\mu x} = (TGx - \mu Tx)' * (TGx - \mu Tx)$
- 5) $C\mu = \sum_{x=1}^i C_{\mu x} / i$
- 6) $\mu Bx = \mu Tx - \mu$
- 7) $\vec{C}_{\mu Bx} = j_i * (\mu Bx)' * (\mu Bx)$
- 8) $C\mu B = \sum_{x=1}^i C_{\mu Bx} / i$
- 9) [Ev, Eval] = Eig($C_{\mu B} / C_{\mu}$)
- 10) $LDA_{\text{training}} = \text{Ev} \times D1_{m \times n} \rightarrow$ Projection

Testing:

- 7) $LDA_{\text{testing}} = \text{Ev} \times D2_{m \times n} \rightarrow$ Projection

Let i be the number of gases and j be the number of concentrations used in the experiment. The data obtained from gases G_1, G_2, \dots, G_i is stored in classes $T_{G1}, T_{G2}, T_{G3}, \dots, T_{Gi}$. After data extraction the mean of each class T_{G_s} is stored as μ_{T_s} , and the overall mean μ is computed as the average of all μ_{T_s} . The boundary-mean μ_B of each class is computed by subtracting the corresponding mean of the class from the average mean μ to estimate the difference between classes. The normalization of class in LDA is performed by subtracting the data of each class from its corresponding class-

mean and then within-class differences is computed by taking the covariance matrix $C_{\mu i}$ of each normalized class. The average of $C_{\mu i}$ provides the within-class-variance (C_{μ}). The average from the product of covariance of the boundary-mean C_{B_i} along with the feature vector size provide the between-class-scattering ($C_{\mu B}$). Since LDA maximizes the between-class-scattering ($C_{\mu B}$) and minimizes the within-class-variance (C_{μ}), therefore, the Eigenvector is computed on the matrix obtained after dividing $C_{\mu B}$ by C_{μ} .

C. Decision Tree (DT) Classifier

The DT is a supervised learning approach and is used in this research because the feature reduction approach is independent of the classifier, therefore a simple but identical classifier is enough to analyze the performance of PCA and LDA based feature reduction. The reason of using DT for classification is because of its implementation simplicity and uniform behavior [30]. CV is used to determine the appropriate component in both reduction approaches which provides the best classification result using DT. The whole process of CV is carried out using MATLAB based simulations and the best scenarios obtained from these simulations are implemented and examined on hardware. It should be noted that the hardware is used only for the testing purpose and therefore the final DT for the best scenarios of PCA/LDA is generated using the entire data [31].

The DT requires a number of predictors which defines the number of variables given to the DT as input. However, the DT can take all or few of them as a selected predictors for tree formation. The designed DT consists of three major parameters which include the decision node (DN), the tree leaves and the tree depth. The DT formation starts from single root DN and expands to further DN in each step until a point is achieved after which no further DN is connected. The point after which no DN is connected and no expansion is possible is referred to as classification point. Thus, the maximum number of steps required to reach the final classification point is used to determine the tree depth. Moreover, the branches which have classification points are termed as tree leaves.

IV. SOFTWARE AND HARDWARE IMPLEMENTATION RESULTS

The training and testing algorithms for both LDA and PCA, along with the DT are first implemented using MATLAB and then the testing part is implemented on hardware. The mean values obtained after performing CV on each feature reduction approach are used to examine various parameters for both PCA and LDA which are required to make a decision about the best classification results. These parameters are used for performing hardware implementation.

The software and hardware implementation timing for both PCA and LDA is also examined to determine the delay caused by them in run-time application. The training task is performed using MATLAB where the simulation timing depends on the processor's speed and varies with the number of active tasks currently handled by the processor. To minimize the effect of this dependency, both PCA and LDA

run on a single program with different variables so that the environment remains the same for both algorithms. Moreover, the implementation is initialized multiple times for the overall sample data and the average is taken as the final result.

The Xilinx Zynq SoC which is used for hardware implementation combines a dual core ARM Cortex-A9 processor, which is the central part of the processing system (PS), with a Xilinx 7 series FPGA, which is the programmable logic (PL). The first step for the hardware implementation of PCA, LDA and DT is to develop the corresponding IP-Cores that will be implemented on the PL using Vivado HLS. Once the IPs are tested and validated within Vivado HLS, they are exported to IP Catalog, and then Vivado IP Integrator is used to create the needed hardware block design that will contain various IPs including the one developed using Vivado HLS. The final step in the hardware implementation is to export the hardware design to Xilinx software development kit (SDK). The hardware implementation design flow is shown in Fig. 6 (a), the steps are as follows:

1. IP design using Vivado HLS

Vivado HLS is used to create the corresponding register transfer level (RTL) design of the system's algorithm described in C. In this step a function named "Predict" is designed, which consists of C source code corresponding to the gas identification system. The input of the function is a vector of 16 floating-point elements in the case of the 4x4 sensor array and 7 floating-point elements in the case of the Figaro sensors while the output is a single integer value. In the case of 4x4 array sensor, the integer will be "1" for CO, "2" for C₂H₆O and "3" for H₂. Similarly, in the case of Figaro sensors the integer value will be "1" for C₃H₈, "2" for CO₂, "3" for H₂ and "4" for NH₃. The vector of means and required eigenvectors are declared and initialized within the function, they are needed for normalization and projection purpose respectively. It is worth mentioning that in the case of PCA both means and eigenvectors are used while in the case of LDA only the eigenvectors are required. The implementation of the DT models generated during the training phase take the form of multiple if-else statements in C. A second C file is needed for testing since in Vivado HLS the test bench is also written in C. The C test bench takes the form of the main C function that will execute the "predict" function and self-check the results. Vivado HLS support different optimization directives, among which loop unrolling, array partitioning and pipelining are applied to test the performance of both algorithms under different hardware implementation approaches. The "Unroll Loop" directive, when applied to a loop in the program, allows iterations of a given loop to be executed in parallel having dedicated hardware resources for each iteration. The second directive, which is "Array Partition" is used to consider a given array in the program as

multiple entities where each entity is having its own data ports compared to one array entity which has limited data ports for data transfer. The last "Pipeline" directive is applied to the top level function of the system to allow pipelining of all instructions and sub function existing inside. Details related to Vivado HLS tool are provided in [32].

2. Implementation on Zynq platform

In this step a hardware block design is developed to allow a program to run on the PS to be able to send and receive data to and from the PL as well as to manage the HLS IP-Cores implemented on the PL. The hardware design is created in Vivado using IP Integrator, this block design will include the IP designed using Vivado HLS along with the IP corresponding to the Zynq PS. Two extra IPs are added to the design, namely "AXI interconnect" and the "Processor System Reset". The AXI interconnect is used as an interface between an AXI memory mapped master device which is the PS in our case, with a memory mapped slave device which is the AXI-Lite compatible HLS core called "Predict" in this solution. The other IP which is the "Processor System reset" provides a customized reset for the entire system, including the PS, the AXI interconnect core and the Predict core from HLS. Details about Vivado IP Integrator can be found in [33]. Fig. 6 (b) shows the hardware block design and how different blocks are connected. It is worth mentioning that the "ZYNQ7 Processing System" IP is not implemented on the PS, it is used for configuration purpose only, the IP is used for example to set the frequency of the PL at 122 MHz, while the HLS IP Core in Fig. 6 (b) is the one that represents our system, it contains either PCA+DT or LDA+DT depending on the scenarios.

3. Software Design Using Software Development Kit (SDK)

A program to get the basic setting and the initialization of the platform, including the universal asynchronous receiver/transmitter (UART) to print results in the terminal is created using SDK and executed on the PS. The C source code of the program is then modified to read/write data from/to the HLS core implemented on the PL. The communication with the hardware present in the PL is performed by calling some read and write data functions that exist in the driver files which were automatically created and exported for various OS including Linux and the lightweight Standalone OS [34]. The hardware results are visualized in terms of resources usage and execution time for the implementation of the EN system on the Zynq-ZC702 prototyping board using different scenarios.

A. Results from 4x4 Sensor Array

The classification results for both PCA and LDA based feature reduction algorithm are shown in Table II. These results are the mean value obtained after 5-fold CV.

TABLE II
CLASSIFICATION RESULTS OBTAINED WITH AND WITHOUT PCA/LDA USING 4X4 SENSOR ARRAY

| Properties | Raw-Data | LDA | | PCA | | | | | |
|--------------------|--------------------|------|------|------|------|------|------|------|------|
| | | 1-DF | 2-DF | 1-PC | 2-PC | 3-PC | 4-PC | 5-PC | |
| Classification (%) | Steady state (SS) | 76.7 | 70.0 | 95.0 | 46.7 | 73.3 | 80.0 | 86.7 | 91.7 |
| | Delta (Δ) | 66.7 | 76.7 | 93.3 | 43.3 | 68.3 | 75.0 | 91.7 | 91.7 |

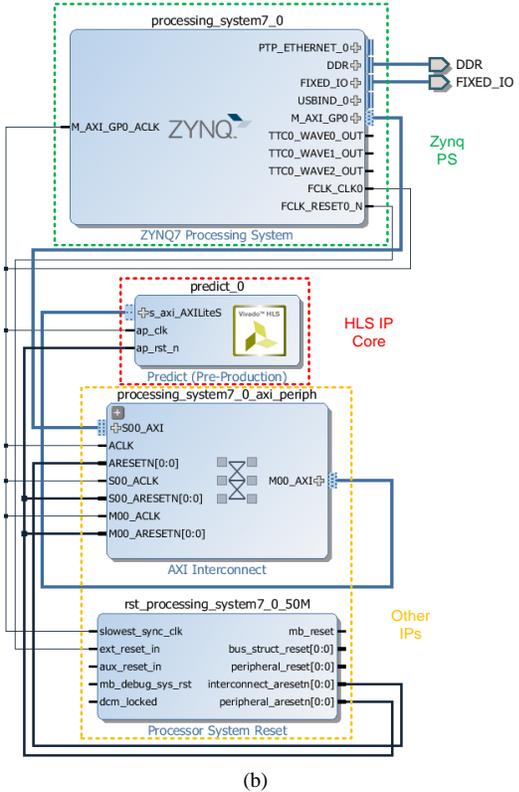
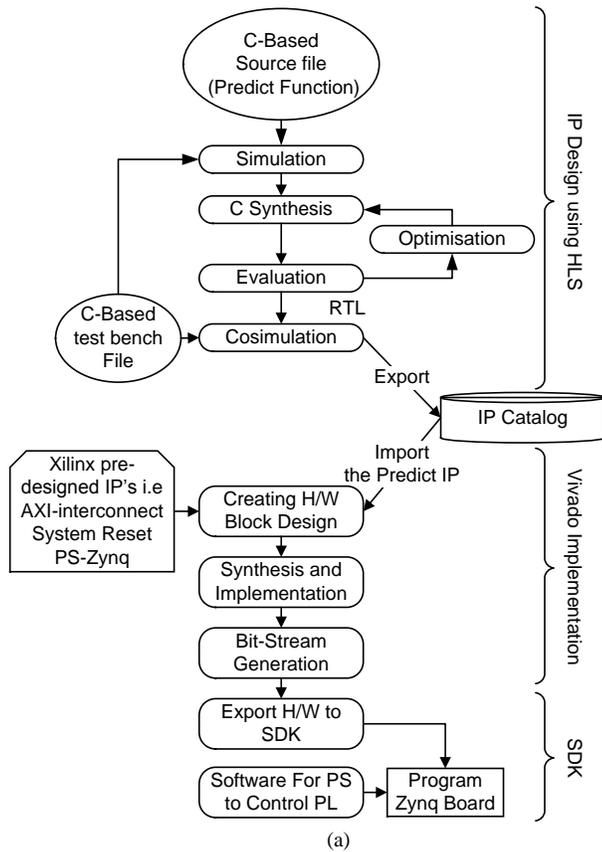


Fig. 6. Hardware Implementation (a) Flow chart (b) Hardware block design

The average classification accuracy of all folds after PCA and LDA based feature reduction are graphically presented using mean and SD values and shown in Fig. 7. It can be seen

from the graph that in case of PCA the best performance of the classifier is obtained using five principal components, after which the performance either becomes constant or degrade a little. However, in case of LDA the best result is obtained using two components. Moreover, the computed values of CoV for each corresponding mean and SD value shows that, in case of PCA with SS of data, the smallest value is obtained with five principal components in which the corresponding CoV is found to be 0.11, which drops to 0.07 when using the Δ . In case of LDA the smallest value of COV is 0.04 to 0.07 at 2-DF with SS and Δ , respectively. However, it can be noted that the values of SD are small in most cases of PCA and LDA with SS and Δ which shows the validity of the computed mean values because the classification at each fold is very near to the final mean-classification. It can also be observed from Table II that the classification obtained from the SS of raw-data without performing any feature reduction algorithm is 76.7% which is reduced to 66.7% for the Δ . However, the first five PC with an accuracy of 91.7%, for SS, provide 15% better classification. While in case of LDA, the DT provides 95.0% identification with only first two DF which is 18.3% better than raw-data and 3.3% better than PCA with 5 PC. It should be noticed that due to the small data size, LDA performs 3.3% better than PCA. However, with the increasing data size the difference in performance will also increase. Moreover, the increased data size will also increase the possibility of outliers to which specially PCA is very sensitive. In order to determine the best feature reduction algorithm for hardware implementation, the complexity of DT and the implementation timing is also an important factor for consideration. It has been observed that 2-DF with two decision nodes and three leaves offers a simpler tree structure than 5-PC as shown in Table III. Moreover, the software DT formation time and the testing time obtained in MATLAB for both algorithms is almost equal, while 2-DF requires 151 ms more pre-processing time than 5-PC which make it almost 7 times slower. It is worth mentioning that in the case of PCA with 5-PC, the 1st, 2nd, 3rd, 4th and 5th PC are given to the DT training phase as predictors. However, only the 1st, 2nd, 4th and 5th are selected by the final generated tree.

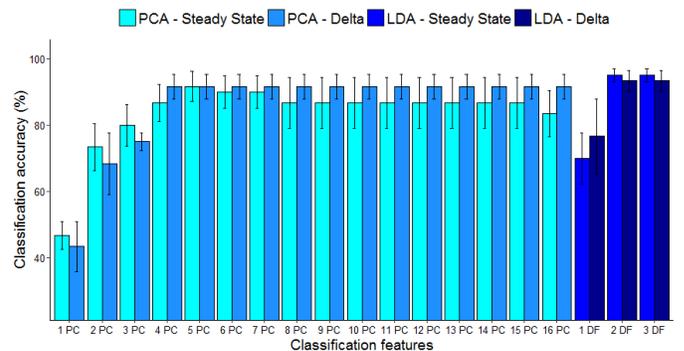


Fig. 7. Mean and SD obtained after CV for PCA and LDA on SS and Delta using 4x4 sensor array.

The data distribution after performing the 2-DF is shown using scatter-plot in Fig. 8 (a). The solid lines in the plot are

showing the values of DN required for the classification of the gases. Data distribution after performing the 5-PC is shown in Fig. 8 (b). It is worth mentioning that the data distribution in this case is presented in six 2-dimensional figures because it is difficult to visualize four components in the same plot. The x-y scaling is same in rows and columns, respectively, therefore it is only shown at the beginning of each row and column. It can be observed from the plots that the data for each gas is widely separated from each other which is the best scenario for classification.

TABLE III
PARAMETERS AND SOFTWARE EXECUTION TIME REQUIRED FOR FINAL DECISION TREE USING 4X4 SENSOR ARRAY

| | | 2-DF | 5-PC |
|------------------------------------|-----------------------------------|-------------|-------------|
| Classification Accuracy | | 95.0 | 91.7 |
| No. of predictors | | 2 | 5 |
| Decision Tree (SS) | No. of selected predictors | 2 | 4 |
| | No. of Decision nodes | 2 | 6 |
| | No. of Tree leaves | 3 | 7 |
| | Tree depth | 2 | 4 |
| Software Execution Time (s) | Pre-Processing | 0.177 | 0.026 |
| | DT formation | 0.015 | 0.015 |
| | Testing | 0.034 | 0.037 |

The results for hardware implementation in terms of

TABLE IV
HARDWARE RESOURCES USAGE AND EXECUTION TIME BEFORE AND AFTER OPTIMIZATION USING 4X4 SENSOR ARRAY

| | Without Optimization | | Unroll Loop Optimizations | | Array Partitioning and Pipelining | |
|---|----------------------|------|---------------------------|------|-----------------------------------|-------|
| | 2-DF | 5-PC | 2-DF | 5-PC | 2-DF | 5-PC |
| DSP48E | 10 | 20 | 5 | 20 | 20 | 44 |
| FF | 1209 | 2669 | 958 | 2537 | 3756 | 8479 |
| LUT | 2281 | 5070 | 1732 | 6191 | 5222 | 11084 |
| First output execution time (ns) | 1353 | 2156 | 746 | 795 | 721 | 803 |
| Next outputs execution time (ns) | 1361 | 2164 | 754 | 803 | 65 | 65 |

TABLE V
CLASSIFICATION RESULTS OBTAINED WITH AND WITHOUT PCA/LDA USING 7 FIGARO SENSORS

| Properties | Raw-Data | LDA | | PCA | | | |
|---------------------------|------------------------------------|------|------------|------|------|------------|------|
| | | 1-DF | 2-DF | 1-PC | 2-PC | 3-PC | |
| Classification (%) | Steady State (SS) | 100 | 100 | 100 | 70.5 | 100 | 100 |
| | Delta (Δ) | 100 | 89 | 100 | 96.5 | 99.5 | 99.5 |

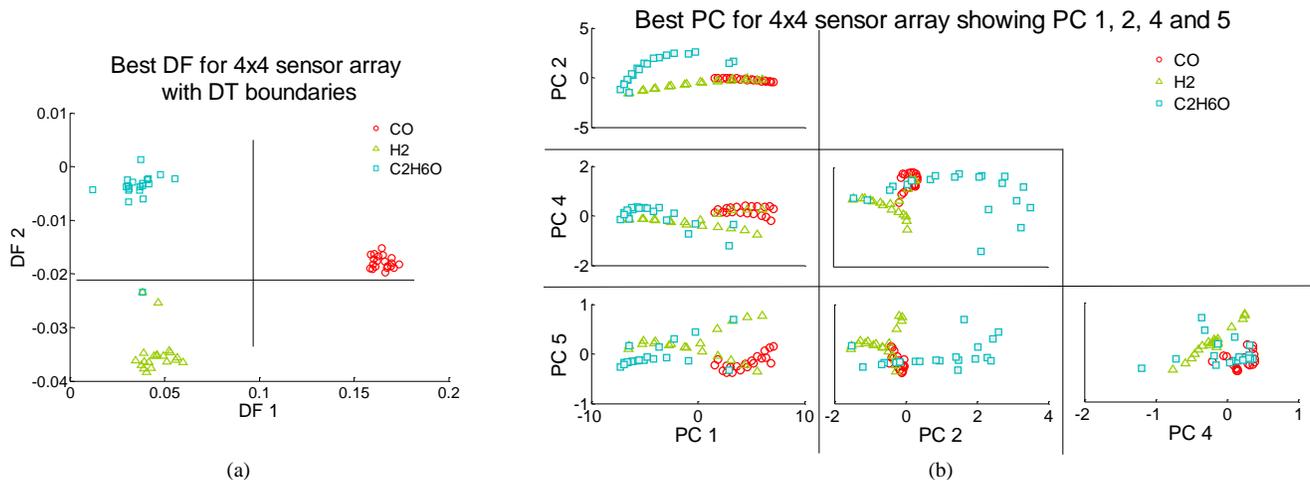


Fig. 8. Data distribution for (a) 2-DF, (b) 5-PC (4 selected), with 4x4 sensor array

resource usage and execution time with and without performing any optimization approach are shown in Table IV. The resources used are related to the number of digital signal processing (DSP) blocks, flip-flops (FF) and lookup-tables (LUT). It has been observed that in all cases of hardware implementation LDA requires an average of 50% less resources and is 10% faster than PCA. Therefore, in terms of both software and hardware implementation LDA is performing better than PCA. The fact that LDA is outperforming PCA in the hardware implementation can be explained by many reasons. Firstly, PCA requires a higher number of components than LDA to provide the best results, 5 and 2 respectively, which in turn increases the size of vectors and matrices. Secondly, for testing purpose PCA requires mean-values for performing the normalization before multiplication with the Eigenvectors while in LDA the mean is only computed in the training phase. During the testing, only the Eigenvector multiplication is performed. The last reason is the size of the tree (DT) which is smaller for the two DFs if compared with the five PCs one.

B. Results from Figaro Sensors

The results from Figaro sensors, as shown in Table V, reveals that the DT can classify 100% of gases from original data without performing feature reduction. However, in this case the objective of feature reduction is to reduce the data size for processing rather than improving the classification. Therefore, a feature reduction algorithm is applied and it is found that DT with first discriminant component and the first two principal components detect 100% of gases with both properties of SS and Δ . The mean and SD values for both PCA and LDA using SS and Δ are presented graphically in Fig. 9(a). Since both PCA and LDA are providing 100% classification which makes it difficult to conclude the best appropriate approach, further simulations are performed. The data used for CV are reduced to 25%, 50% and 75% to check which approach provides 100% classification with minimum sample data. It should be noted that the random sampling approach is adopted for the reduced data, such that the class proportions are not maintained. Moreover, both PCA and LDA approaches are tested with identical samples to obtain and analyze a valid comparison between them. The new simulations are performed using SS since it is the one that shows 100% accuracy almost everywhere. The mean and SD values obtained after CV for SS of each sample size are shown graphically in Fig.9 (b). It is observed that the SD in the case of Figaro sensor is also small which validates the mean values obtained from CV. Also, the LDA based feature reduction reaches 100% classification accuracy by using 50% of the sample data which has not been observed in the case of PCA where 100% classification is obtained only when the entire dataset is used, as shown in Fig. 9(b). It can also be visualized that PCA seems to perform better for smaller data size, which may be due to the fact that PCA is sensitive toward the outliers and the randomly sample data increases the possibilities of outliers because the class proportions are not maintained.

Moreover, the final DT obtained using the overall sample data requires a similar number of nodes and tree leaves for both 1-DF and 2-PC as shown in Table VI. The selection of

only these two cases is due to the fact that these two are providing the best classification rate for Figaro sensors with overall sample data. The result shows that both 1-DF and 2-PC require same number of decision nodes and tree leaves. However, the depth of DT for 1-DF is one step more than 2-PC which is because a single predictor is used for classification. Moreover, both algorithms require almost similar time for DT formation and testing. However, 1-DF requires 188 ms more pre-processing time than 2-PC which make it 8 times slower. This is due to the fact that in case of PCA, the pre-processing only refer to mean-normalization of the data, whereas for LDA, within-class-difference and between-class-scattering needs to be computed along with the normalization of each class. The data distribution after performing the 2-PC and 1-DF is shown using scatter-plot in Fig. 10 (a) and (b), respectively. Similar to 4x4 sensor array case, the solid lines in the plot show the values of DN required for the classification of the gases. It can be observed from the plot that the data for each gas is widely separated from each other which is the best scenario for classification.

The results for hardware implementation with and without optimization are shown in Table VII. It has been found that in the case of Figaro sensors as well, the LDA requires an average of 50% less resources than the PCA and is 10% faster. This is similar as in the case of 4x4 array sensor, therefore it is concluded that for the best classification scenario LDA cause less hardware overhead than PCA.

TABLE VI
PARAMETERS AND SOFTWARE EXECUTION TIME REQUIRED FOR DECISION TREE USING 7 FIGARO SENSORS

| | | 1-DF | 2-PC |
|------------------------------------|-----------------------------------|------------|------------|
| Classification Accuracy | | 100 | 100 |
| Decision Tree (SS) | No. of predictors | 1 | 2 |
| | No. of selected predictors | 1 | 2 |
| | No. of Decision nodes | 3 | 3 |
| | No. of Tree leaves | 4 | 4 |
| | Tree depth | 4 | 3 |
| Software Execution Time (s) | Pre-Processing | 0.215 | 0.027 |
| | DT formation | 0.016 | 0.017 |
| | Testing | 0.030 | 0.033 |

TABLE VII
HARDWARE RESOURCES USAGE AND EXECUTION TIME AFTER AND BEFORE OPTIMIZATION USING 7 FIGARO SENSORS

| | Without Optimization | | Unroll Loop Optimizations | | Array Partitioning and Pipelining | |
|---|----------------------|------|---------------------------|------|-----------------------------------|------|
| | 1-DF | 2-PC | 1-DF | 2-PC | 1-DF | 2-PC |
| DSP48E | 5 | 10 | 5 | 10 | 10 | 24 |
| FF | 763 | 1640 | 755 | 1533 | 1458 | 3526 |
| LUT | 1471 | 2939 | 1629 | 3213 | 2521 | 5483 |
| First output execution time (ns) | 615 | 959 | 360 | 410 | 360 | 410 |
| Next outputs execution time (ns) | 623 | 967 | 369 | 418 | 32 | 32 |

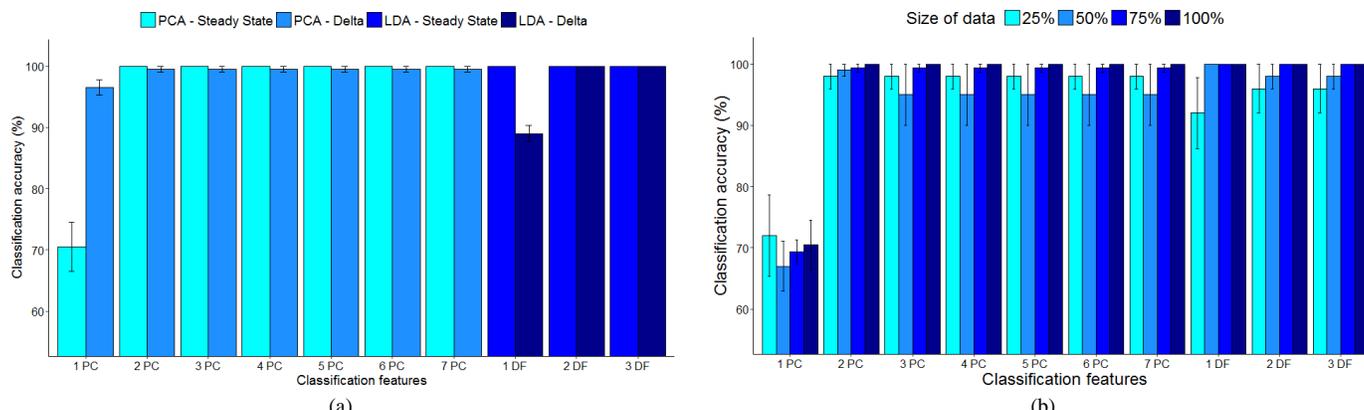


Fig. 9. Mean and SD obtained after CV for PCA and LDA on (a) SS and Delta, (b) different sample size using 7 Figaro sensors

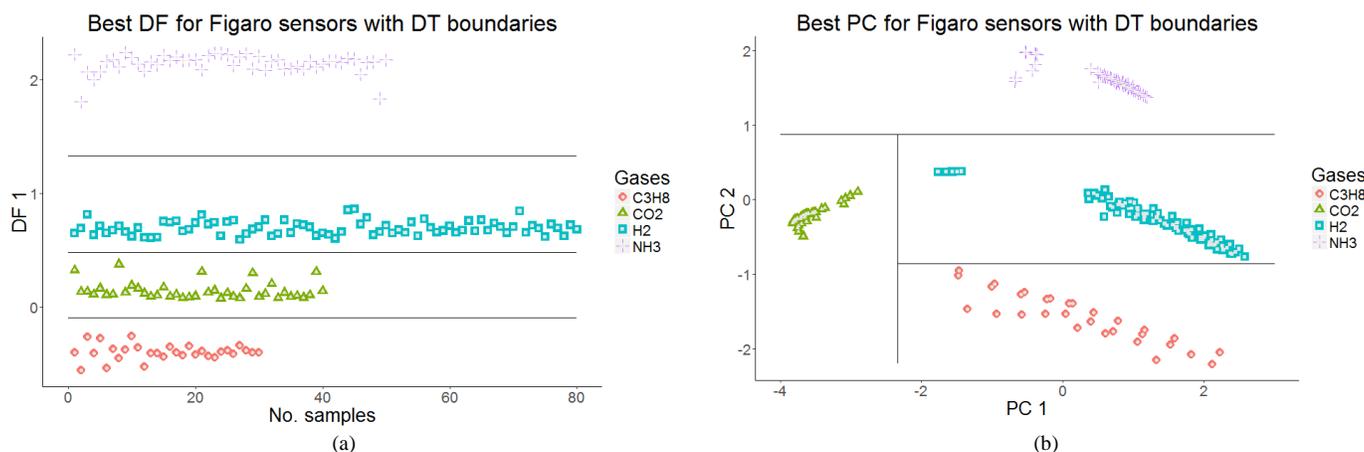


Fig. 10. Data distribution along with the DN values for (a) 1-DF, (b) 2-PC, with 7 Figaro sensors.

V. CONCLUSION

A comparative analysis for PCA and LDA based feature reduction is carried out using MATLAB for software implementation and the Zynq SoC for hardware implementation. Two types of gas sensors have been used, an in-house fabricated 4x4 sensor array and 7 commercial Figaro sensors. In total, six different gas mixtures are used. A 5-fold CV approach has been adopted to statistically analyze and verify the obtained classification results. Moreover, two different properties, SS and Δ , obtained from the data are used for analysis. A DT based classifier is used to obtain the performance matrix after each reduction approach. The results from 4x4 sensor array reveal that the classification results obtained from DT using 2-DF are 3.3% more accurate than the 5-PC. The small classification difference is due to the small data size. However, it should be noted that with the increasing data size the difference in performance between PCA and LDA may also increase. Similarly, the classification of gas is improved to 18.3% by using LDA as compared to the classification performed on raw-data.

In case of Figaro sensors, DT can classify 100% of gases using raw-data obtained directly from the sensors. However, the goal of feature reduction is to minimize the data size without losing the most important information. Therefore, feature reduction approaches are applied and it is found that

with the reduced feature size, 2-PC and 1-DF also provides 100% classification with identical tree parameters. In order to conclude the one best feature reduction approach for the case of Figaro sensors, simulations have been performed using different proportions of the collected data. The obtained results reveal that LDA reaches 100% classification accuracy with only 50% of the collected sample data, whereas for PCA 100% classification accuracy can only be reached using the entire collected dataset. The goal of this research is to analyze the appropriate feature reduction approach between PCA and LDA which can be advantageous for software and hardware implementation. It is concluded that for the given sample data LDA with fewer components provide better classification than PCA. Similarly, in terms of hardware implementation, LDA based feature reduction outperforms PCA with approximately 50% less resources along with the 10% speed efficiency. Furthermore, the PCA requires large memory size to store the Eigen vectors along with corresponding means whereas in case of LDA the memory is required to only store the Eigen vectors. Therefore, the low memory requirement along with the corresponding computation time, makes LDA more suitable for hardware implementation than PCA.

REFERENCES

- [1] V. V. Sysoev, B. K. Button, K. Wepsiec, S. Dmitriev & A. Kolmakov, "Toward the nanoscopic 'electronic nose': hydrogen vs carbon monoxide discrimination with an array of individual metal oxide nano- and mesowire sensors", *Nano letters*, 2006, 6, (8), pp. 1584-1588.
- [2] V. V. Sysoev, J. Goschnick, T. Schneider, E. Strelcov & A. Kolmakov, "A gradient microarray electronic nose based on percolating SnO₂ nanowire sensing elements", *Nano letters*, 2007, 7, (10), pp. 3182-3188.
- [3] B. Guo, A. Bermak, P.C. Chan & G. Z. Yan, "An integrated surface micromachined convex microhotplate structure for tin oxide gas sensor array", *IEEE Sensors Journal*, 2007, 7, (12), pp. 1720-1726.
- [4] A. Vergara, R. Calavia, R. M. Vázquez, A. Mozalev, A. Abdelghani, R. Huerta, E. H. Hines & E. Llobet, "Multifrequency interrogation of nanostructured gas sensor arrays: a tool for analyzing response kinetics", *Analytical chemistry*, 2012, 84, (17), pp. 7502-7510.
- [5] M. Shi, A. Bermak, S. B. Belhouari, & P. C. Chan, "Gas identification based on committee machine for microelectronic gas sensor", *IEEE Transactions on Instrumentation and Measurement*, 2006, 55, (5), pp. 1786-1793.
- [6] D. Molnar, P. Heszler, R. Mingesz, Z. Gingl, A. Kukovecz, Z. Konya et al., "Increasing chemical selectivity of carbon nanotube-based sensors by fluctuation-enhanced sensing", *Fluctuation and Noise Letters*, 2010, 9, (03), pp. 277-287.
- [7] E. Llobet, R. Ionescu, S. Al-Khalifa, J. Brezmes, X. Vilanova, X. Correig, B. Nicolae & J. W. Gardner, "Multicomponent gas mixture analysis using a single tin oxide sensor and dynamic pattern recognition", *IEEE Sensors Journal*, 2001, 1, (3), pp. 207-213.
- [8] L. B. Kish, R. Vajtai & C. G. Granqvist, "Extracting information from noise spectra of chemical sensors: single sensor electronic noses and tongues", *Sensors and Actuators B: Chemical*, 2000, 71,(1), pp. 55-59.
- [9] R. Gutierrez-Osuna, "Pattern analysis for machine olfaction: a review", *IEEE Sensors Journal*, 2002, 2, (3), pp. 189-202.
- [10] H. T. Nagle, R. Gutierrez-Osuna, B. G. Kermani and S. S. Schiffman. "Environmental monitoring", *Handbook of machine olfaction: Electronic nose technology*, T. C. Pearce, S. S. Schiffman, H. T. Nagle and J. W. Gardner (Eds.), Wiley-VCH, pp 457-440, 2002.
- [11] T. Hastie, R. Tibshirani, J. Friedman & J. Franklin, "The elements of statistical learning: data mining, inference and prediction", *The Mathematical Intelligencer*, 2005, 27, (2), pp. 83-85.
- [12] S. Z. Li, X. Lu, X. W. Hou, X. Peng & Q. Cheng, "Learning multiview face subspaces and facial pose estimation using independent component analysis", *IEEE Transactions on Image Processing*, 2005, 14, (6), pp. 705-712.
- [13] N. P. Chandrasiri, M. C. Park, T. Naemura & H. Harashima, "Personal facial expression space based on multidimensional scaling for the recognition improvement", *Proc. of the Fifth International Symposium on Signal Processing and Its Applications*, 1999, 2, pp. 943-946.
- [14] P. Honeine, "Online kernel principal component analysis: A reduced-order model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34, (9), pp. 1814-1826.
- [15] A. Sophian, G. Y. Tian, D. Taylor & J. Rudlin, "A feature extraction technique based on principal component analysis for pulsed Eddy current NDT", *NDT & E International*, 2003, 36, (1), pp. 37-41.
- [16] V. Monasterio, P. Laguna & J. P. Martinez, "Multilead analysis of T-wave alternans in the ECG using principal component analysis", *IEEE Transactions on Biomedical Engineering*, 2009, 56, (7), pp. 1880-1890.
- [17] L. Qingzheng and A. Bermak. "A low-power hardware-friendly binary decision tree classifier for gas identification", *Journal of Low Power Electronics and Applications*, 2011, 1(1), pp. 45-58.
- [18] T. K. Kim & J. Kittler, "Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image", *IEEE Transactions on Analysis and Machine Intelligence*, 2005, 27, (3), pp. 318-327.
- [19] S. Chakrabarti, S. Roy & M. V. Soundalgekar, "Fast and accurate text classification via multiple linear discriminant projections", *The VLDB Journal*, 2003, 12, (2), pp. 170-185.
- [20] S. Dudoit, J. Fridlyand & T. P. Speed, "Comparison of discrimination methods for the classification of tumors using gene expression data", *Journal of the American statistical association*, 2002, 97, (457), pp. 77-87.
- [21] Q. Gao, L. Zhang, D. Zhang & H. Xu, "Independent components extraction from image matrix", *Pattern Recognition Letters*, 2010, 31,(3), pp. 171-178.
- [22] Z. Ankara, T. Kammerer, A. Gramm and A. Scutze. "Low power virtual sensor array based on a micromachined gas sensor for fast discrimination between H₂, CO and relative humidity", *Sensors and Actuators B: Chemical*, 2004, 100, (1), pp. 240-245.
- [23] 'Inc Xilinx. Zynq-7000 All Programmable SoC: Technical Reference Manual', http://www.xilinx.com/support/documentation/user_guides/ug585-Zynq-7000-TRM.pdf, UG585, v1.8.1, accessed 02 Feb. 2015.
- [24] 'Figaro Engineering Inc.', <http://www.figarosensor.com/>, accessed 22 April 2015.
- [25] T. C. Pearce, S. S. Schiffman, H. T. Nagle & J. W. Gardner (Eds.), "Handbook of machine olfaction: electronic nose technology", John Wiley & Sons, pp. 133-155.
- [26] W. Song & X. Shaowei, "Robust PCA based on neural networks", *Proc. of the 36th IEEE Conference on Decision and Control*, 1997, 1, pp. 503-508.
- [27] K. T. Ng, F. Boussaid & A. Bermak, "A CMOS single-chip gas recognition circuit for metal oxide gas sensor arrays", *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2011, 58, (7), pp. 1569-1580.
- [28] M. E. Tipping & C. M. Bishop, "Probabilistic principal component analysis", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 1999, 61, (3), pp. 611-622.
- [29] E. K. Tang, P. N. Suganthan, X. Yao & A. K. Qin, "Linear dimensionality reduction using relevance weighted LDA", *Pattern recognition*, 2005, 38, (4), pp. 485-493.
- [30] R. Chutia & M. Bhuyan, "Online concentration independent feature dimension reduction of metal oxide gas sensor based E-Nose", *International Conference on Advances in Engineering, Science and Management (ICAESM)*, 2012, pp. 247-250.
- [31] G. C. Cawley and L. T. Nicola. "On over-fitting in model selection and subsequent selection bias in performance evaluation", *The Journal of Machine Learning Research*, 2010, 11, pp. 2079-2107.
- [32] 'Inc Xilinx. Vivado Design Suite User Guide: High-Level Synthesis', UG902,v2014.4. http://www.xilinx.com/support/documentation/sw_manuals/xilinx2014_3/ug902-vivado-high-level-synthesis.pdf, accessed 30 March 2015.
- [33] 'Inc Xilinx. Vivado Design Suite User Guide: Designing IP Subsystems Using IP Integrator', UG994, v2014.4. http://www.xilinx.com/support/documentation/sw_manuals/xilinx2014_3/ug994-vivado-ip-subsystems.pdf, accessed 30 March 2015.
- [34] 'Inc Xilinx. Vivado Design Suite User Guide: Embedded Processor Hardware Design.', UG898, v2014.4. http://www.xilinx.com/support/documentation/sw_manuals/xilinx2014_3/ug898-vivado-embedded-design.pdf, accessed 30 March 2015.



Muhammad Ali Akbar received the B.E. degree in Electronics with distinction from NED University of Eng. & Tech., Pakistan, and the M.Sc. degree in computer engineering from Chosun University, Korea.

He is currently a Research Assistant in the dept. of Computer Eng. Qatar university, where he is conducting research on automated structural health monitoring system. His research interest includes artificial intelligence, reconfigurable computing, embedded systems design and reliability.



Amine Ait Si Ali received a BEng degree from The University of Science and Technology Houari Boumediene, Algiers, Algeria, in 2009 and an MSc degree in embedded intelligent systems from the University of Hertfordshire, Hatfield, UK, in 2012. He is currently a full time research assistant in the School of

Engineering, Qatar University, Doha, Qatar and a final year distance PhD student in the school of Engineering and Computing, University

of the West of Scotland, Paisley, UK. His research interests are mainly in custom computing using FPGAs, heterogeneous embedded systems, machine learning and connected.



Prof. Abbas Amira (M'01–SM'07) received the Ph.D. degree in computer engineering from Queen's University Belfast, Belfast, U.K., in 2001. He held academic and consultancy positions in U.K. and overseas, including the positions of Professor of Computer Engineering with Qatar University, Doha, Qatar and Full Professor of Visual Communications and a Leader of the Visual Communication Cluster with the University of the West of Scotland, Paisley, U.K. He held other academic positions with the University of Ulster, U.K., Qatar University, Brunel University, U.K., and Queen's University Belfast. He was a Visiting Professor with the University of Lorraine, Nancy, France and Tun Hussein Onn University of Malaysia, Johor Bahru, Malaysia. He is a regular referee for many national and international funding bodies, including the Engineering and Physical Research Council and the Qatar National Research Fund, and he sits on the international advisory boards for some international research centers. He has authored around 300 publications in reconfigurable computing, and image and signal processing. His current research interests include embedded systems, high-performance reconfigurable computing, image and video processing, multiresolution analysis, biometrics, and connected health applications.

Prof. Amira is a fellow of the Institution of Engineering and Technology, and the Higher Education Academy, and a Senior Member of the Association for Computing Machinery. He has been awarded a number of grants from the government and industry. He has been invited to give talks, short courses, and tutorials at universities and international conferences, and is the Chair and Program Committee Member for a number of conferences. He was one of the tutorials presenters at the International Conference on Image Processing in 2009, the Chair of the Embedded Computer Vision Workshop (ECVW) in 2011, the Program Chair of ECVW in 2010, the Program CoChair of the International Conference on Microelectronics (ICM) in 2012, the DELTA Conference in 2008, the Irish Machine Vision and Image Processing Conference in 2005, and the General Co-Chair of ICM in 2014.



F. Bensaali (S'03–M'06–SM'15) obtained a Dipl-Ing (MEng) in electronics from University of Constantine and a PhD in computer science from Queen's University, Belfast. He is currently an Associate Professor in Electrical Engineering at Qatar University. Dr Bensaali took other academic positions at the University of Hertfordshire-UK and Queen's University Belfast-UK. His research interests are mainly in design and implementation of digital image and signal processing algorithms, custom computing using FPGAs, hardware/software co-design and system on chip. He is an HEA Associate and IEEE Senior member.



Mohieddine Benammar (M'00–SM'05) received the Maîtrise (B.Sc.) degree in electrical engineering and the DEA (M.Sc.) degree in automation and control from the École Normale Supérieure de l'Enseignement Technique de Tunis (currently École Supérieure des Sciences et Techniques de Tunis), University of Tunis II, Tunisia, in 1985 and 1987, respectively. He was awarded a British Council scholarship enabling him to join the Gas Sensors Research Group at the School of Engineering, Middlesex University, London, U.K., where he was awarded the Ph.D. degree in Sensors and Instrumentation in 1991. Subsequently, he spent three years as Post-Doctoral Research Fellow with Middlesex University. Then he joined the University of Tunis and University of Monastir as Assistant Professor. He moved to Qatar University, Doha, Qatar, in 1997. He is currently with the College of Engineering at Qatar University where he holds the position of Full Professor of electronics and instrumentation. He was head of department of Electrical Engineering at Qatar University from 2009 to 2014, and Co-Chair of Industrial Electronics for six years, a position funded by LNG producers RasGas at Qatar University from 2008 to 2014. His fields of interest include general electronic instrumentation, solid-state sensors, bioinstrumentation, and engineering education.



Muhammad Hassan received his MPhil degree in Electronics and Computer Engineering from Hong Kong University of Science and Technology (HKUST) in 2012. He is currently pursuing for the PhD degree in the same university. His interests include development of pattern recognition algorithms for odor discrimination and development of energy aware platform for structural health monitoring applications.



Prof. Amine Bermak received the Masters and PhD degrees, both in electrical and electronic engineering (microelectronics and Microsystems), from Paul Sabatier University, Toulouse, France in 1994 and 1998, respectively. Prof. Bermak has received Six distinguished awards, including the "Best University Design Contest Award" at ASP Design Automation Conference (DAC) in Macau 2016; "Best student paper award" at IEEE International Symposium on Circuits and systems ISCAS 2010; the 2004 "IEEE Chester Sall Award" from IEEE Consumer electronics society; the IEEE Service Award from IEEE Computer Society and the "Best Paper Award" at the 2005 International Workshop on System-On-Chip for Real-Time Applications. He has published over 250 articles in journals, book chapters and conference proceedings and designed over 30 chips. He has graduated 14 PhD and 16 MPhil students. He has served on the editorial board of IEEE Transactions on Very Large Scale Integration (VLSI) Systems and the Sensors Journal. He is also currently serving on the editorial board of IEEE Transactions on Biomedical Circuits and Systems and the IEEE Transactions on Electron Devices (TED). Prof. Bermak is a Fellow of IEEE and IEEE distinguished Lecturer.