IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 17, 2020, accepted March 30, 2020, date of publication April 13, 2020, date of current version April 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2987756

An End to End Indoor Air Monitoring System Based on Machine Learning and SENSIPLUS Platform

MARIO MOLINARA[1], (Member, IEEE), MARCO FERDINANDI[1,2,*], (Member, IEEE), GIANNI CERRO[3], (Member, IEEE), LUIGI FERRIGNO[1], (Member, IEEE), AND ETTORE MASSERA[4]

[1]Department of Electrical and Information Engineering, University of Cassino and Southern Lazio, 03042 Cassino, Italy
[2]Sensichips s.r.l., 04011 Aprilia, Italy
[3]Department of Medicine and Health Sciences, University of Molise, 86100 Campobasso, Italy
[4]ENEA, 80055 Portici, Italy

Corresponding author: Mario Molinara (m.molinara@unicas.it)

*Marco Ferdinandi contributed to the article until October 31, 2019.

The work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under Agreement 787128, and in part by the Minister for Education, University and Research (MIUR) Law, Department of Excellence under Grant 232/216.

ABSTRACT In the framework of indoor air monitoring, this paper proposes an Internet of Things ready solution to detect and classify contaminants. It is based on a compact and low–power integrated system including both sensing and processing capabilities. The sensing is composed of a sensor array on which electrical impedance measurements are performed through a microchip, named SENSIPLUS, while the processing phase is mainly based on Machine Learning techniques, embedded in a low power and low resources micro controller unit, for classification purposes. An extensive experimental campaign on different contaminants has been carried out and raw sensor data have been processed through a lightweight Multi Layer Perceptron for embedded implementation. More complex and computationally costly Deep Learning techniques, as Convolutional Neural Network and Long Short Term Memory, have been adopted as a reference for the validation of Multi Layer Perceptron performance. Results prove good classification capabilities, obtaining an accuracy greater than 75% in average. The obtained results, jointly with the reduced computational costs of the solution, highlight that this proposal is a proof of concept for a pervasive IoT air monitoring system.

INDEX TERMS Contaminant detection, air monitoring, sensor networks, neural networks, deep learning, IoT.

I. INTRODUCTION

Air monitoring is a topic for which last few years have witnessed a deep increase of interest in many different fields. Lots of efforts have been addressed particularly for applications regarding the citizens' health and safety care, as cities' pollution control [1]–[3]. Pollution monitoring for the detection and identification of contaminants are only some but surely among the principal applications driving the development of new ubiquitous and low–cost air monitoring sensing technologies [4]. Nowadays, bulky and costly devices requiring the employment of professional technicians are currently adopted for air monitoring related tasks. Indeed, in the pollution control application, static or mobile (mounted on vehicles) cumbersome stations are used for air quality monitoring in urban centers. Their high costs are usually the main reason for a very sparse or, in same cases (e.g. poorer countries), non–existent air quality control. For this reason, a system that is, at the same time, reliable from the sensing point of view and allowing a dense diffusion with low costs is the goal of many research activities [5], [6]. The interest of the scientific community is focused on either the sensing technology or the data analysis. The first issue is addressed with the development of miniaturized sensors able to respond to contaminants with the same performance level than bulky and costly systems, while the second is pursued through the

The associate editor coordinating the review of this manuscript and approving it for publication was M. Anwar Hossain.

rise of novel algorithms and processing techniques able to retrieve information about contaminants starting from raw measurement data. In this sense, the authors propose an integrated system, able to optimize both the sensing and processing, responding to desirable requirements such as low cost, integration, portability, light computational burden, good sensitivity and classification capability. In detail, stemming from the authors' experience in gas recognition [7], water analysis [8], [9], an indoor air monitoring system, based on a compact and low cost sensing technology and Artificial Intelligence (AI) techniques for the detection and classification of air contaminants, is proposed. The system development starts from the assumption, well known in scientific literature on this field, that a single sensor is seldom sensitive to more contaminants and therefore the best solution is the employment of a sensor array, as implemented in this work. Furthermore, having more than one sensor causes the rapid growth of the amount of data to be processed to obtain classification. Such issue has encouraged the researchers to focus on ad–hoc computing techniques, often based on decentralized architectures [10]. According to this trend, the authors have looked for techniques able to manage and fully exploit large data sizes, generally available under Machine Learning (ML) based approaches. Aiming to minimize the computational burden, a lightweight Multi Layer Perceptron (MLP) solution, as in [11], is proposed and compared with more complex Deep Learning (DL) architectures [12] as Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM). Three are the main novelties of the paper: a) the use of an innovative measuring chip (i.e. the SENSIPLUS) capable of embedding up to 6 air quality sensors, directly inserted on the chip surface; b) the embedding of sensing, measuring and classification on reduced resources and costs platforms according to IoT and Edge Computing paradigms; c) the comparison of DL and ML classification performance in terms of accuracy and computational burden.

The paper is organized as follows. Section II gives an overview about the state of art of current detection and classification advances in research field; in section III a comprehensive description of the system is given, both in terms of sensing and data processing. In this section a detailed description of the sensors peculiarities, data acquisition and preprocessing, and neural network training is provided. Obtained results are given in section IV and a discussion regarding the achieved goals in terms of application requirements is provided. Finally, conclusions follow in section V.

II. STATE OF THE ART

A. SENSING ISSUES

The scientific sensing scenario has highlighted various technological limitations as the low sensitivity and selectivity or the environmental and atmospheric conditions dependencies [13] of miniaturized, low cost and smart solutions. In [5], Castel *et al.* analyze low cost commercial platforms' performance versus CEN (European Committee for Standardization) reference instruments, highlighting the fact that despite of a lower accuracy, stability and selectivity they provide a useful added value represented by the possibility to perform data aggregation. Spatial analysis as mapping and gradient could allow to evaluate the pollutants sources. The well-known chemical micro-sensors limitations are deeply faced in the scientific literature through various techniques. A review of methodologies to improve the performance of chemical sensors in different tasks as classification, regression and clustering is provided in [14].

B. PROCESSING ISSUES

Different Solutions for classification are available in literature. Among the others, here we mention: i) the Decision tree induction, ii) the rule-based methods, iii) the support vector machine and iv) those based on neural networks. Starting from some considerations reported in [15], i.e. algorithms based on neural networks show better classification accuracy and considering the noisy of the measured data that may prevent the reliability of algorithm based on thresholds [16], the authors have paid their attention on the iv) category of classification algorithm. AI and more in particular ML techniques have been adopted in order to find hidden correlations among sensors array responses to chemical substances. In [17], Esposito *et al.* face the problem of on field calibration of a low cost technology through a Dynamic Neural Network approach. In [18] and [19] the authors exploit ML techniques for gas recognition and concentration estimation, respectively. In [20] the authors describe a carbon monoxide and methane quantification system based on a sensor array and an Artificial Neural Network (ANN). The technological revolution experienced in the ML field with the born of DL has furthered its adoption in an increasing number of different applications outperforming other classical ML techniques. Its capability to automatically extract data features and to profit from large amount of data have been key elements for its adoption also in sensors based applications.In [21] and [22], proofs of how DL overcomes other classical ML techniques in air quality monitoring are provided. In [23] Peng *et al.* propose a gas classification system based on CNN comparing its performance against the ones obtained with MLP and Support Vector Machine (SVM) architectures. CNNs represent a breakthrough experienced with the born of DL for different fields, as image or time series analysis as highlighted in [24] by Bengio *et al.* An important DL architecture, well suited for time series analysis is represented by LSTM neural network [25]. It is a special kind of Recurrent Neural Networks (RNN), which are networks with loops capable to maintain the information. The computed output at each step is inputted for the next step, providing the possibility to find out and correlate time dependencies. The greater complexity of the LSTM internal units allows to counteract the long short term dependency, which represents one of the main RNN limits [26]. In [27] and [28], deep RNN and LSTM are applied to different scenarios with important data time dependencies.

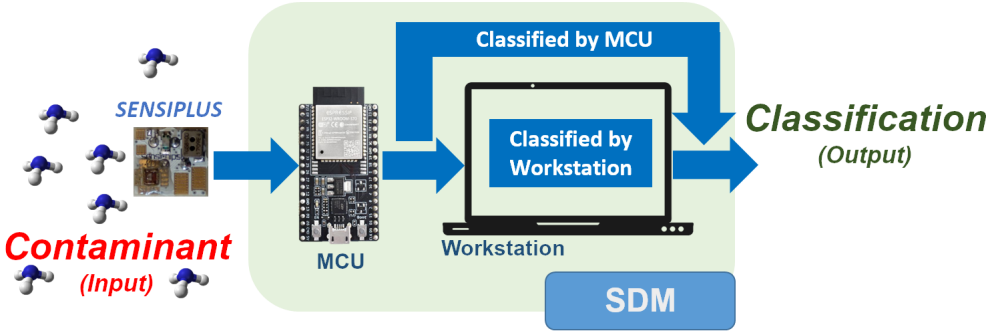

FIGURE 1. The proposed integrated system. SDM stands for SENSIPLUS Deep Machine.

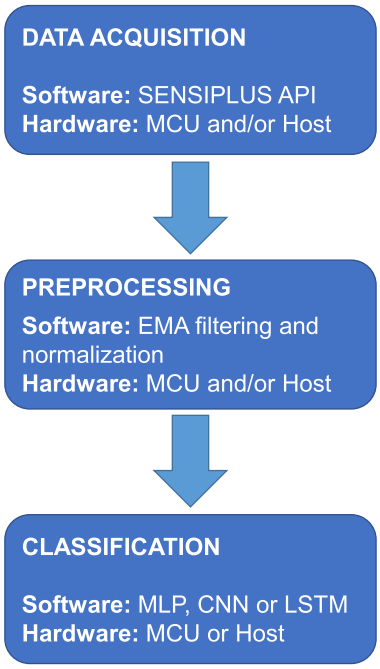

FIGURE 2. The SDM block diagram.

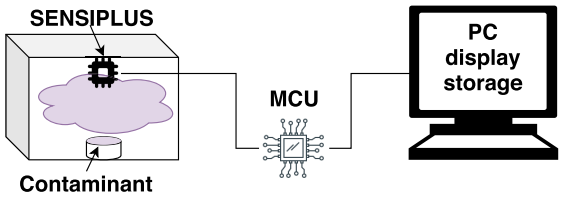

FIGURE 3. Measurement setup.

III. PROPOSED INTEGRATED SYSTEM

A. OVERVIEW

The proposed integrated system is shown in Figure 1 and mainly composed of:

- SENSIPLUS Chip (henceforth SPC): it is a microelectronic measurement device endowed with on-chip sensing capabilities developed by Sensichips s.r.l. [29] and the Department of Information Engineering of the University of Pisa. Endowed with a versatile analog front end and different internal and external ports, it allows to perform electrical impedance measurements both on internal and external sensors. It has been already adopted in other works, as in [7], [30], [31].
- SENSIPLUS Deep Machine (SDM): a hardware/software module for data acquisition, processing and analysis. The block diagram, depicted in Figure 2, shows the logical operation flow and highlights the software and hardware components exploited for each task. Data Acquisition task is carried out through the SPC API, which is a software library, developed either in Java or C programming language, runnable on both Micro Controller Unit (MCU) and Linux/Windows/Android hosts. Acquired data are then preprocessed through an ad–hoc developed module, which is described in detail in section III-B.4. As for the API, depending on the application requirements, it can be executed on a MCU and/or a host. Finally, classification can be performed through one of the adopted ML techniques (MLP, CNN or LSTM). The ML technique can be run on a MCU or on a more computationally endowed device as a PC, depending on which one is chosen.

B. DATA ACQUISITION AND PREPROCESSING

The experimental campaign has been carried out in order to acquire a set of raw measurements used as the basis for ML training phases (further described in III-C).

1) MEASUREMENT SETUP

The adopted measurement setup, shown in Figure 3, is composed of:

- a Personal Computer (PC) running a proprietary (ad hoc developed) JAVA software for measurement storage and displaying;
- a MCU running SPC API and transferring acquired values to the PC through USB. The MCU is connected to SPC with the proprietary one wire serial communication protocol, namely SENSIBUS;
- a SPC plugged on the cable disposed on the internal surface of a transparent glass box and endowed with 3 sensors;

- a little bowl containing the contaminant in the liquid state used to submit the gas to sensors through the evaporation process.

The adopted experimental setup is aimed to emulate the real scenario which correspond to a common indoor environment where low-concentrations of contaminants can be found. Tens of ppm is the concentrations expected in reference application and, for this reason, a small scale emulation of the same conditions has been pursued. Furthermore, as a proof of concept research activity, the efforts have been focused on the system sensitivity and classification capability.

As for the sensing technology, 3 different sensors, whose operating principle is provided in III-B.2, have been used:

- The internal sensor based on aluminum oxide interdigitated electrodes, namely ONCHIP_ALUMINUM _OXIDE;
- An external commercial capacitive humidity sensor, namely OFFCHIP_HUMIDITY;
- An external sensor based on gold interdigitated electrodes functionalized with graphene as sensing material, namely OFFCHIP_GRAPHENE.

The electrical impedance measuring capability of the SPC has been exploited for the measurement acquisition phase. A preliminary sensitivity analysis has allowed to optimize the measuring settings (e.g. sinusoidal stimulus frequency and amplitude) and to select a specific electrical quantity (e.g. resistance, capacitance, conductance etc.), both aiming to maximize the sensor sensitivity. Furthermore, the aforementioned analysis phase has brought to witness, as highlighted by the scientific literature, the sensors' dependence on the environmental condition, in terms of response's slope and amplitude. Moreover, in order to follow the continual variation of the environment (temperature and humidity), a reference quantity has been computed through Exponential Moving Average (EMA) (III-B.4) for each sensor.

2) NOTES ABOUT ADOPTED SENSORS

As listed in III-B.1, the adopted sensors in the experimental campaign are characterized by different peculiarities.

The ONCHIP_ALUMINUM_OXIDE is the SPC built-in sensor based on aluminum oxide interdigitated electrodes. It is a generic sensor capable to fastly react to volatile compounds. Water vapour and contaminants molecules depositing among the interdigitated tracks and affecting the aluminum oxide cause a variation of the electrical properties, allowing their detection. The OFFCHIP_HUMIDITY is a simple capacitive sensor manufactured by IST ([32]). The operating principle is based on the dielectric constant variation of the polymer used as sensing material. The typical capacitance value (23 °C and 30% RH) is 140 ± 40 pF, measured in the frequency range 1 kHz – 100 kHz.

Finally, the adoption of OFFCHIP_GRAPHENE has been enabled by the research activity carried out at ENEA Portici research center [33]. Here we report a summary of the properties and its fabrication for sensing purposes. Graphite flakes were obtained [34] from NGS Naturgraphit GmbHWinner Company (Leinburg-Germany). Iso-Propyl Alcohol (IPA) was purchased from Carlo Erba. All aqueous solutions were prepared with ultrapure water from Type1 Ultrapure Milli-Q system (Millipore). Pristine graphene was synthesized from natural graphite powder by a Liquid Phase Exfoliation (LPE) method. The process is a sonication–assisted exfoliation of graphite flakes in a hydro–alcoholic solution. Specifically, 80 mg graphite flakes were dispersed into 80 ml of a water/IPA mixture (7:1 v/v). The dispersion was sonicated in a low–power bath (around 30 W) for 48 h. Afterwards, unexfoliated graphitic crystals were separated from the dispersion by centrifugation at 500 rpm for 45 min obtaining a black, homogeneous suspension of few–layer graphene at concentration of 0.1 mg/ml. A flake of graphene, based on a few layers of carbon sheets, is an almost two-dimensional material and the chemical change on its surface can drastically change the electric transport on the entire flake. Graphene flakes are sensitive to oxidizing and reducing gases. A particular chemical and electrical affinity for nitrogen dioxide is experienced. It can significantly modify the resistivity of a film of graphene flakes already to a few tens of ppb in the humid air. A graphene flake film can be easily dispensed onto an interdigitated electrode by dropcasting a few tens of microliters of the graphene suspension.

3) MEASUREMENT PROCEDURE

hlThe experimental campaign has been carried out according to a systematic procedure collecting data with a substance at a time. All sensors' responses are acquired with an acquisition rate, empirically chosen according to the observed phenomenon velocity, of 0.5 S/s performing a measurement from the three sensors at each acquisition time–step. The measurements have been achieved pursuing the following phases:

- Air exposure for 120 seconds;
- Chemical substance introduction inside the glass box for 600 seconds;
- Air exposure for further 120 seconds.

As regards the initial air exposure, it has been empirically chosen after a preliminary analysis to let the sensors reach the steady state. The fixed time interval of 600 seconds, used for the chemicals evaporation phase, has been chosen according to an experimental evaluation of the sensors response time and to meet the fastly responsive application requirement. Finally, the last 120 seconds of clean air exposure have been used to analyze the sensors recovery capability. With the selected sampling rate, 420 ($= 60 + 300 + 60$) samples are collected for each experiment. 10 repetitions have been conducted for each substance to analyze the measurement repeatability and enhance the ML system generalization capability among different environmental conditions. The following dangerous substances, commonly found in indoor environments, have been adopted: acetone, alcohol, ammonia, bleach. Furthermore, to estimate the sensing system measurement background, further data acquisitions have been carried out submitting water vapour and clean air to

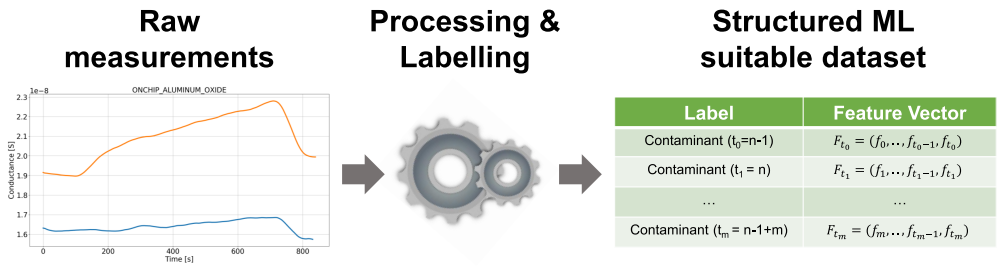

FIGURE 4. ML suitable Dataset preparation schema. Parameter n and m correspond to the time-window size of each sample and the dataset size, respectively.

the sensors. In this way, whatever classification technique is involved, it has been trained to distinguish a chemical substance from air and water vapour to avoid classification errors.

4) PREPROCESSING

The huge amount of raw measurements acquired through the setup described in the previous section has required the development of an *ad–hoc* preprocessing software module, whose operation principle is shown in Figure 4. It deals with a series of preliminary operations on the sensors responses and the generation of a labelled dataset suited for the successive ML training phase. The developed software allows to adapt the raw measurements to any of the adopted ML techniques' requirements.

In this module, the main operations are: the evaluation of an environmental reference, normalization and labelling. The first one is accomplished by employing an Exponential Moving Average (EMA) filtering stage to acquired samples. Such operation is addressed to face the problem of the sensors' baseline continuous variation, mainly caused by the environmental conditions fluctuation and sensors' drift. The EMA filtering has been computed according to equation 1, where s_t is the 3–dimensional measurement vector at time instant t, $\alpha = 10^{-4}$ the degree of weighting decrease and e_t the environmental reference vector computed at the same time instant.

$$e_t = \alpha s_t + (1 - \alpha)e_{t-1} \tag{1}$$

The α coefficient has been chosen to have a high relevance to initial samples, during the clean air sensors exposition. Although this approach suffers long term dependency on the response to contaminants exposition, the proposed system is not deeply influenced since it is designed to provide early detection and recognition. As for the normalization, the ratio between each sensor response and its e_t parameter is computed, as in equation (2).

$$f_t^i = s_t^i / e_t^i, i = 1, 2, 3 \; with f_t = (f_t^1, f_t^2, f_t^3) \tag{2}$$

Finally, the labelling operation is performed on the normalization output, assigning the substance identifier to each sample selected through a specific time window and step. A time window of parametric size (n) can be generated for each sample, providing in such a way a bi-dimensional feature vector ($F_t = f_{t-n+1}, .., f_{t-1}, f_t$).

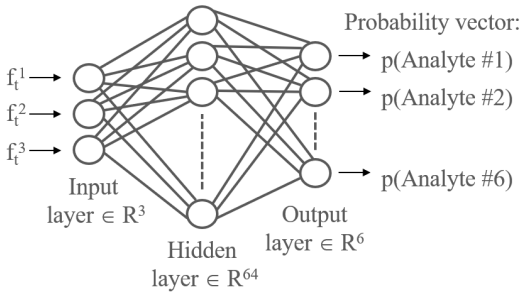

FIGURE 5. Implemented MLP_3 architecture.

C. ADOPTED MACHINE LEARNING ARCHITECTURES

As stated in the introduction, the goal is to provide the user with a classification system able to perform most operations on the MCU. Therefore, we implemented and performed classification by means of a lightweight MLP. Nevertheless, to compare our results with heavier, more complex and powerful classifiers, we also tested our dataset with two other ML and particularly DL approaches, namely the CNN and LSTM. In the following, a more detailed description of the designed architectures is presented. A preliminary tuning phase of the networks' hyper-parameters has been executed for each architecture. Setting parameters, defining the networks' structure (e.g. number of layers and internal neurons) and the ones determining the way they are trained (e.g. learning rate, drop–out and batch normalization), have been selected in this stage.

1) MULTI LAYER PERCEPTRON

The designed MLP, as depicted in Figure 5, is characterized by 3 input neurons, 64 neurons in the hidden layer and 6 output neurons. The Rectified Linear Unit (ReLU) activation function has been applied on the output of the hidden layer and the softmax function has been chosen for the output layer. The 3–dimensional feature vector, computed according to equation 2, is used as input for this network. In such a way, no time latency and memory usage are required since a classification output is generated for each acquisition time–step. Since, as described in section III-C.2, for the DL architectures a 2–dimensional feature vector (F_t) has been used as input,

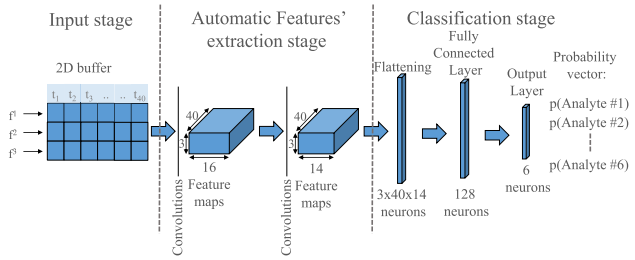

FIGURE 6. Developed Convolutional Neural Network (CNN) architecture.

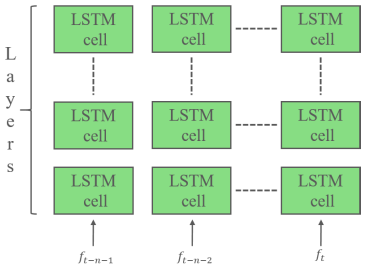

FIGURE 7. Unrolled LSTM architecture.

a further implementation of the MLP model has been carried out for a fairer comparison. Here, the same F_t, flattened in a mono–dimensional feature vector, has been used as input. For this reason, a wide input layer with 120 neurons has been adopted, while the number of neurons in the hidden layer has been preserved. In section IV-B, results corresponding to both implementations, from now on referred as MLP and MLP_{120}, are presented.

2) CONVOLUTIONAL NEURAL NETWORK

The developed CNN, as shown in Figure 6, is characterized by the following stages:

- Input: a bi–dimensional feature vector ($F_t = f_{t-n+1}, .., f_{t-1}, f_t$) is used as input, where n is equal to 40.
- Automatic features' extraction: 2 consecutive Convolutional Layers (CLs) are used in this stage for automatic features' extraction. Both the CLs execute the same operations: zero padding of size 1, convolution, batch normalization and ReLU activation function. Convolutional kernels sizes are $(16 * 3 * 3)$ and $(14 * 16 * 3 * 3)$ for CL_1 and CL_2, respectively. In such a way, the output of this stage is a 3–dimensional matrix with size $(14 * 3 * 40)$.
- Classification: the computed feature maps in the previous stage are flattened in a mono-dimensional vector and used as input for the Fully Connected Layers (FCLs). The latter is composed of 128 neurons (hidden layer) that are fully connected to the 6 neurons of the output layer.

3) LONG SHORT TERM MEMORY (LSTM) NEURAL NETWORK

A multivariate 8–layer LSTM neural network has been adopted and the same input as for the CNN has been used for this model. Figure 7 shows an unrolled view of the proposed architecture, where each LSTM cell is composed of 240 internal units that contains the trainable parameters.

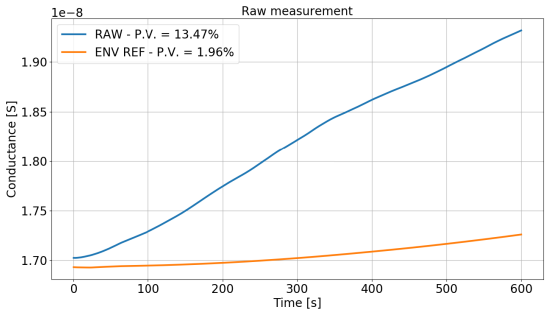

FIGURE 8. Raw sensor response to ammonia. ONCHIP_ALUMINUM_OXIDE sensor response is depicted here as example.

4) NETWORKS TRAINING

As reported in section B, for each substance and for each of the experiments about 420 measurements have been collected. A subset of 300 values has been considered for the network training (the samples obtained in step two of the procedure, after the chemical substance introduction). In this way, a dataset of about $(6 * 10 * 300) = 18000$ samples has been collected. A 10–fold cross–validation has been performed in order to have a statistical analysis of the results over the whole dataset. 1800 samples have been used as test set for each fold (an entire experiment for each substance $(6*1*300)$) while the remaining samples have been exploited for training and validation sets (6 experiments for training and 3 for validation with 10800 and 5400 samples respectively). The network training has been carried out using the Cross Entropy as loss function and Adam Optimizer as optimizing algorithm. Furthermore, an early stopping strategy, based on a maximum number of epochs (patience) without improvements on the validation set, has been used to avoid the network overfitting. Training experiments have been carried out through a Linux server characterized by an Intel(R) Xeon(R) E5-2609 v4 (8 cores) CPU and 256 GB RAM. GPU acceleration has been exploited for all the experiments using a NVIDIA® TITAN X Pascal.

IV. RESULT

A. RAW AND pre–processed MEASUREMENTS

Ten different experiments for each substance have been performed according to the previously explained measurement phases (III-B.3). Once the whole raw dataset has been acquired, it has been submitted to the preprocessing module described in section III-B.4. For instance, the EMA filtering output computed on the ONCHIP_ALUMINUM_OXIDE sensor to ammonia is depicted in Figure 8. The blue line represents the raw sensor response and the orange one is the filtered version.

As highlighted in the Figure 8, a 13.47% Percentage Variation (PV) has been obtained with the raw measurement

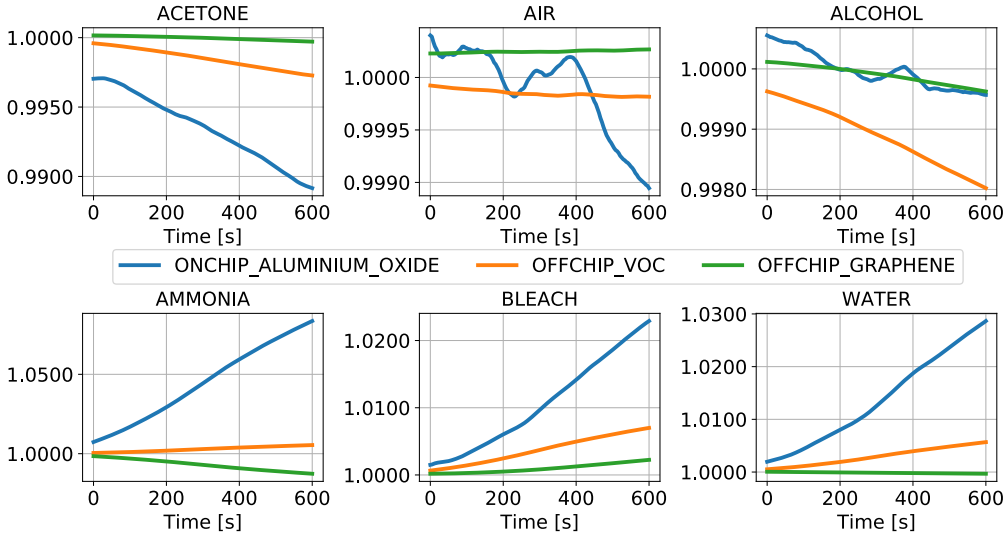

FIGURE 9. Normalized sensors responses to all tested substances.

while a 1.96% PV with the filtered version. The reported PVs prove the negligible effect of the EMA long term dependency problem characterized by a short observation period.

A summarized overview of the normalized sensors responses for each experiment is provided in Figure 9.

The vertical dotted orange lines divide the sensors responses to the various substances, alphabetically ordered in these graphs. All the air exposure phases (initial and final) are removed to focus the attention on the responses to chemicals. Similar behavior characterizes the ONCHIP_ALUMINUM_OXIDE and OFFCHIP_HUMIDITY sensors' responses, confirming their analogous sensing nature, a great complementarity is provided by the OFFCHIP_GRAPHENE sensor. Having an evident sensitivity to only 2 (ammonia and bleach) of the 6 total substances and, more in particular, an opposite trend for them, the OFFCHIP_GRAPHENE sensor is expected to improve the system classification capability.

B. CLASSIFICATION RESULTS

Results obtained with the adopted ML architectures are summarized in Table 1 in terms of global accuracy mean and standard deviation values (evaluated on the 10 folds), while detailed metrics for each class are provided in Tables 2, 3, 4. As for the latters, Accuracy (eqn. 3), Precision (eqn. 4), Recall (eqn. 5) and F1–score (eqn. 6) metrics are shown.

$$\textbf{Accuracy} = \frac{\text{Correctly Classified Samples}}{\text{Total Samples}} \tag{3}$$

$$\textbf{Precision} = \frac{\text{Positive Correctly Classified}}{\text{Total Positive Classified Samples}} \tag{4}$$

$$\textbf{Recall} = \frac{\text{Positive Correctly Classified}}{\text{Total Positive Samples}} \tag{5}$$

$$\textbf{F1–score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

TABLE 1. Mean and accuracy values evaluated for each ML architecture.

Architecture	Mean (μ)	Standard Deviation (σ)
MLP	71.1 %	7.9 %
MLP_{120}	62.4 %	10.8 %
CNN	**75.1** %	**5.6** %
LSTM	69.9 %	12.1 %

TABLE 2. Synthetic performance metrics for MLP architecture: mean value and standard deviation.

Substance		Precision	Recall	F1-score	Accuracy
Acetone	μ	0.50	0.51	0.48	0.51
	σ	0.32	0.35	0.28	0.32
Air	μ	0.53	0.88	0.66	0.90
	σ	0.09	0.19	0.13	0.18
Alcohol	μ	0.52	0.44	0.44	0.44
	σ	0.32	0.35	0.32	0.34
Ammonia	μ	0.96	0.85	0.89	0.84
	σ	0.13	0.09	0.06	0.08
Bleach	μ	0.93	0.84	0.88	0.84
	σ	0.09	0.06	0.06	0.06
Water	μ	0.81	0.76	0.77	0.82
	σ	0.13	0.16	0.13	0.12

As shown, best performance results have been obtained with the CNN but, considering the concept of measurement compatibility, no considerable improvements have been reached with the two DL techniques. A more detailed overview of the obtained results, except for the MLP_{120} which has been used as benchmark, is provided in the form of Confusion Matrices in Figures 10,11 and 12.

The μ and σ values shown in a generic *(i,j)* matrix cell represent the mean and the standard deviation (computed on the 10-fold cross validation testing results) of the samples that belong to class *i* and have been classified as *j*. High accuracies with relative small standard deviations have been obtained for ammonia, bleach and water. A performance decrease is shown by the CNN for air, with respect to MLP and

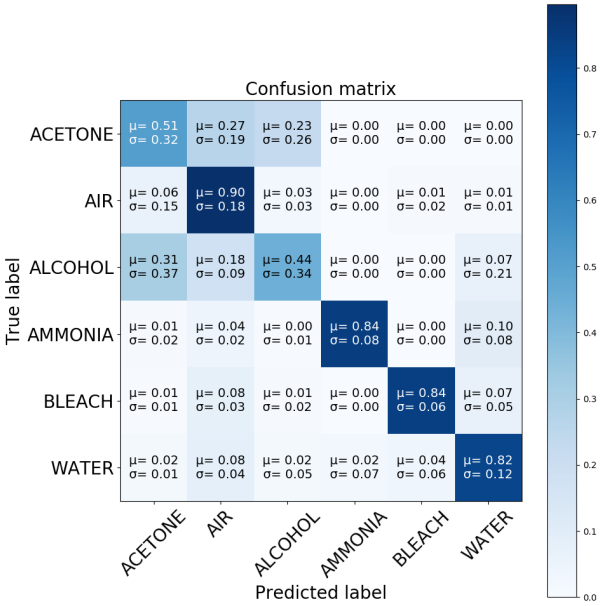

FIGURE 10. MLP Global Confusion Matrix.

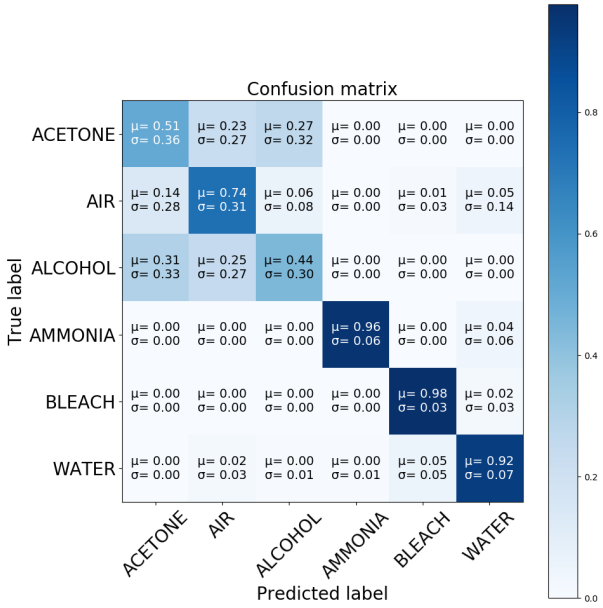

FIGURE 11. CNN Global Confusion Matrix.

LSTM which have obtained 90% and 100% accuracy values, respectively.

Finally, most of confusion regards acetone and alcohol classes: for these contaminants all 3 architectures have shown worst results. For such phenomenon, further considerations are provided in section IV-C.

For the sake of completeness has been evaluated the Receiver Operating Characteristics of the classifiers (see Figures 13–16). As it is possible to see from the figures, the trend of the three curves is equivalent to the ranking of the three classifiers obtained in terms of accuracy, precision, etc. Indeed the three mean curves have an Area Under the

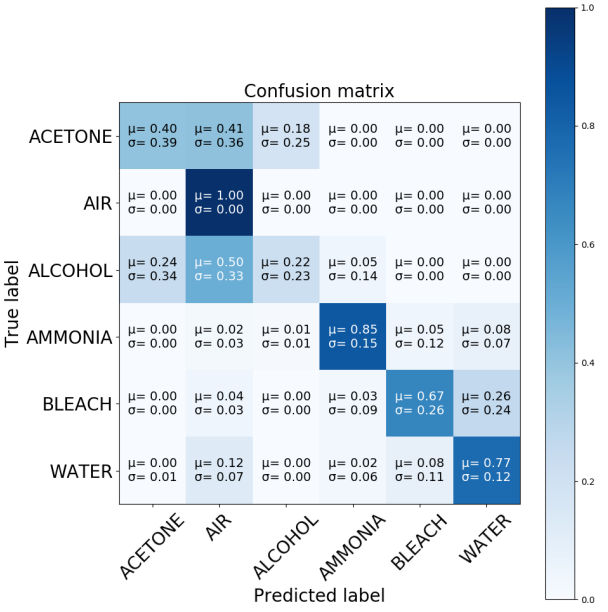

FIGURE 12. LSTM Global Confusion Matrix.

TABLE 3. Synthetic performance metrics for CNN architecture: mean value and standard deviation.

Substance		Precision	Recall	F1-score	Accuracy
Acetone	μ	0.49	0.47	0.43	0.51
	σ	0.32	0.38	0.32	0.36
Air	μ	0.58	0.74	0.62	0.74
	σ	0.19	0.32	0.25	0.31
Alcohol	μ	0.53	0.40	0.40	0.44
	σ	0.38	0.32	0.28	0.30
Ammonia	μ	0.99	0.95	0.97	0.96
	σ	0.03	0.06	0.03	0.06
Bleach	μ	0.90	0.98	0.94	0.98
	σ	0.10	0.03	0.03	0.03
Water	μ	0.86	0.81	0.82	0.92
	σ	0.16	0.16	0.13	0.07

TABLE 4. Synthetic performance metrics for LSTM architecture: mean value and standard deviation.

Substance		Precision	Recall	F1-score	Accuracy
Acetone	μ	0.51	0.41	0.42	0.40
	σ	0.38	0.38	0.38	0.39
Air	μ	0.62	0.89	0.72	1.00
	σ	0.25	0.25	0.25	0.00
Alcohol	μ	0.58	0.51	0.48	0.22
	σ	0.32	0.35	0.32	0.23
Ammonia	μ	0.95	0.87	0.90	0.85
	σ	0.16	0.09	0.09	0.15
Bleach	μ	0.67	0.64	0.64	0.67
	σ	0.35	0.38	0.35	0.26
Water	μ	0.66	0.71	0.67	0.77
	σ	0.28	0.28	0.25	0.12

Curve (AUC) equal to 0.92 for CNN, 0.91 for MLP and 0.87 for LSTM.

C. DISCUSSION

1) FURTHER INVESTIGATIONS ON CLASSIFICATION RESULTS

As introduced in section IV-B, a common trend has been obtained with all ML architectures. Two key observations

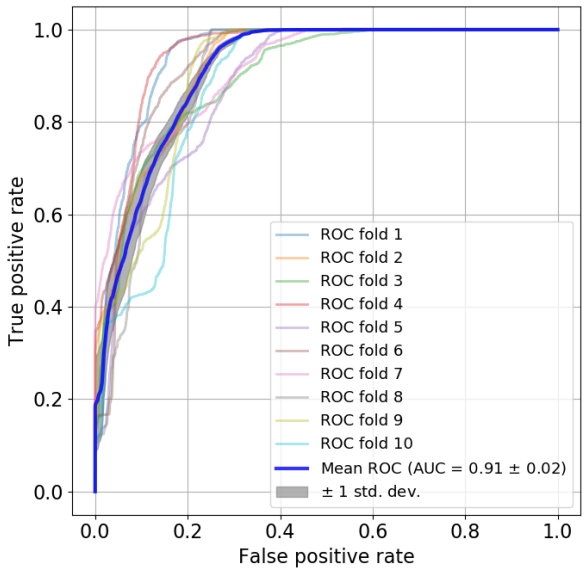

FIGURE 13. The ROC curves for MLP.

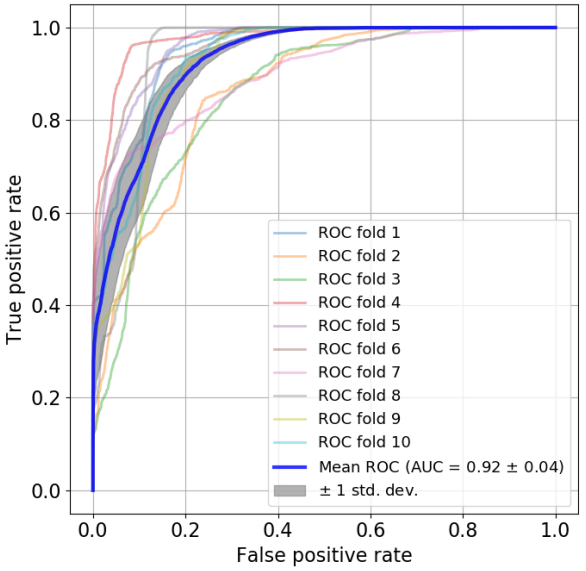

FIGURE 14. The ROC curves for CNN.

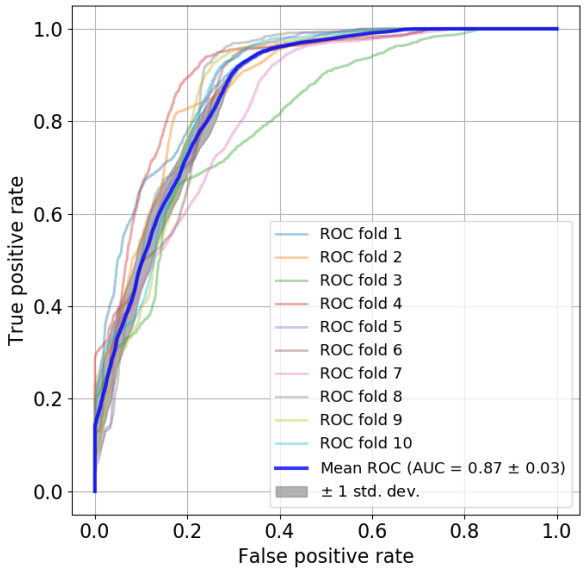

FIGURE 15. The ROC curves for LSTM.

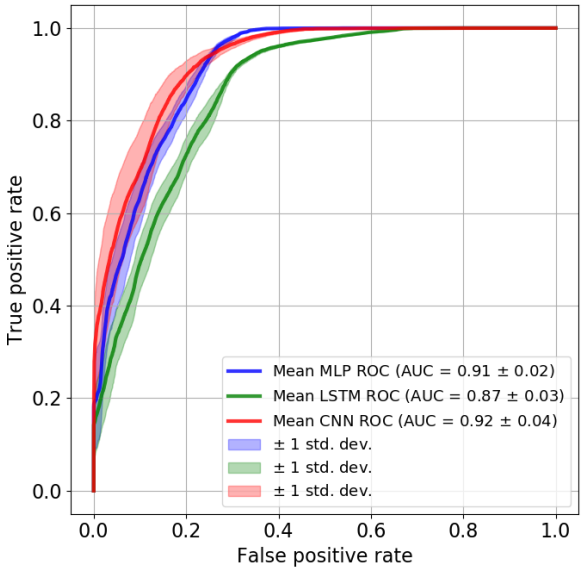

FIGURE 16. The mean ROC curve of the three classifiers with the AUC for CNN, MLP and LSTM respectively of 0.92, 0.91, 0.87.

can be further highlighted: for air, ammonia, bleach and water classes a high accuracy is obtained, meaning that the integrated system has a good sensitivity and recognition capability for these substances. As regards acetone and alcohol, they are both polar substances that can attack oxygen atoms on the film sensitive surface. Furthermore, they both have low boiling points; therefore, they are difficult to be distinguished by the adopted sensor array. For this reason, in Figures IV-C.1–IV-C.1 the classifiers' outputs, obtained with the aggregation of data coming from both acetone and alcohol, are reported. In detail, we have considered samples belonging to these classes labeling them as one only substance, namely acet–alc. With this operations, the element (1,1) of the confusion matrices shows a higher recognition rate, comparable with those belonging to the remaining classes.

2) COMPUTATIONAL ANALYSIS AND MEMORY FOOTPRINT OF MLP AND CNN

An evaluation of the FLoating point OPerationS (FLOPS) and Memory Footprint (MF) in bytes both for MLP and CNN is provided in this section, since they are the lightest and the one providing best accuracy, respectively.

The CNN, as described in section III-C.2, is composed of two convolutional layers (CL1, CL2) as feature extraction stage and a final set of 2 FC layers used for classification, while the MLP contains only the FC layers. A general evaluation of the needed FLOPS for the two convolutional layers and for the FC is provided in equation 7, whose parameters are defined in Table 5. Please note that the MLP is completely

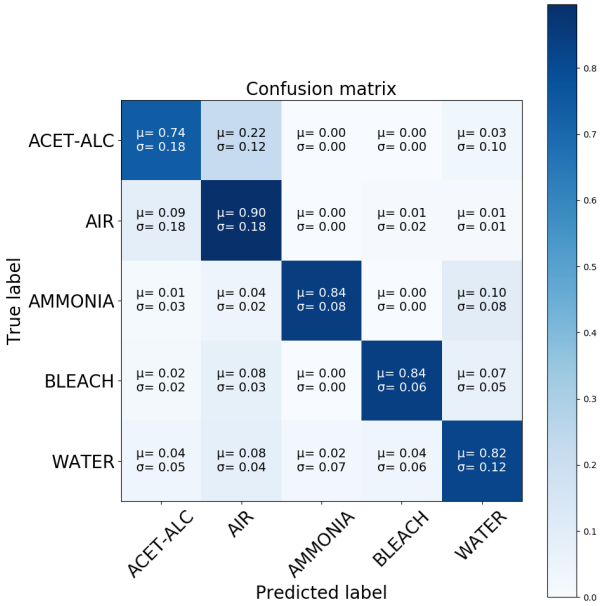

FIGURE 17. MLP Global Confusion Matrix.

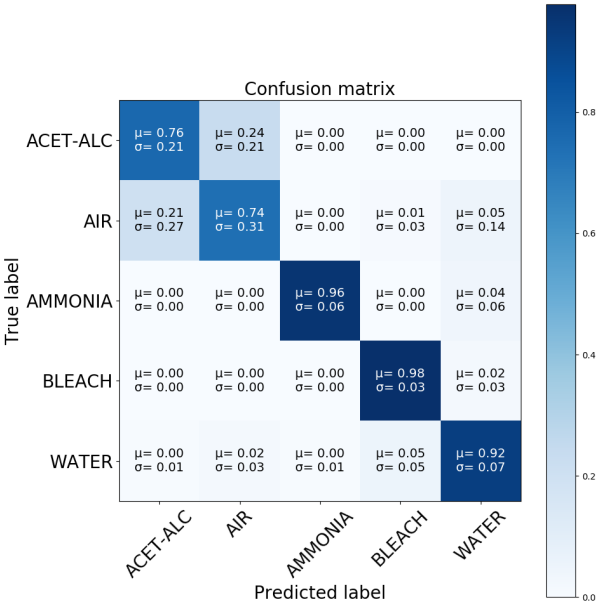

FIGURE 18. CNN Global Confusion Matrix.

characterized by the FC layers.

$$FLOPS = \begin{cases} wsk_1(2k^2+1) & \text{CL1} \\ wsk_2(2k^2k_1+1) & \text{CL2} \\ 2n_1(n_0+n_2) & \text{FC} \end{cases} \quad (7)$$

The overall memory footprint (X) is the sum of two contributions, the first related to the model parameters (X_p), the second related to the run-time (X_r). By considering that on the ESP32 the float type is represented with 4 bytes, in the

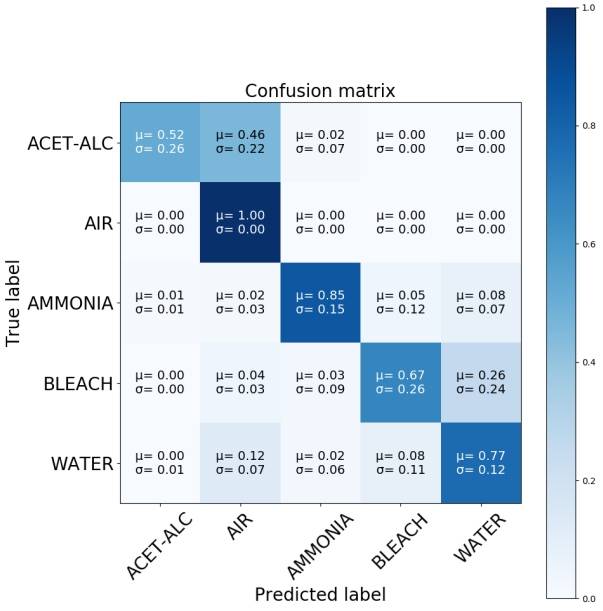

FIGURE 19. LSTM Global Confusion Matrix.

TABLE 5. Adopted notation symbols for computational analysis and memory footprint.

Symbol	Description	Value (MLP)	Value (CNN)
w	timesteps	1	40
s	sensors	3	3
k_1	1st layer kernels	N/A	16
k_2	2nd layer kernels	N/A	14
k	kernel size	N/A	3
n_0	input neurons	3	$w \cdot s \cdot k_2$
n_1	hidden neurons	64	128
n_2	output neurons	6	6

TABLE 6. Evaluated number of floating point operations for MLP and CNN.

Technique	CL1	CL2	FC	Tot.
MLP	N/A	N/A	1152	1152
CNN	36480	485520	462848	984848

equations (8) appear 4 as a multiplicative coefficient.

$$MF = \begin{cases} 4k_1k^2 & \text{CL1}_p \\ 4wsk_1 & \text{CL1}_r \\ 4k^2k_1k_2 & \text{CL2}_p \\ 4wsk_2 & \text{CL2}_r \\ 4[(n_0+1)n_1+(n_1+1)n_2] & \text{FC}_p \\ 4(n_0+n_1+n_2) & \text{FC}_r \end{cases} \quad (8)$$

The obtained results in terms of FLOPS and MF for each layer of the two ML techniques are presented in Table 6 and 7. As expected, MLP is clearly much lighter than CNN because of the lack of the features' extraction layers and the reduced neurons number used as FC input (s vs wsk_2) and hidden layers (64 vs 128).

In the next section detail about the adopted MCU is given. Here can be observed that from Table 7 is evident that

TABLE 7. Evaluated MF in bytes for MLP and CNN both for parameters (p) and runtime (r) variables).

Technique	CL1	CL2	FC	Tot.
MLP_p	N/A	N/A	2584	2584
MLP_r	N/A	N/A	292	292
CNN_p	576	8064	863768	872408
CNN_r	7680	6720	7256	21656

the number of bytes requested by CNN (21656 + 872408) is greater than the pSRAM dimension. The problem has been resolved by using the flash memory for parameters and pSRAM for runtime. Please note that generally speaking, data can be read from flash memory as many times as needed, while most devices are designed and tested for about 100,000 to 1,000,000 write operations.

3) FULL SYSTEM TIME AND POWER ANALYSIS

Both MLP and CNN have been trained on a powerful server, therefore the implementation on the MCU regarded only the classification phase. This means that, once the network weights have been correctly set, they are used for all classification trials. A characterization of execution time for measurement acquisition and classification through MLP or CNN has been performed through an embedded implementation of the SDM on the ESP32 MCU. The adopted MCU has a 32-bit CPU Xtensa dual-core LX6 microprocessor, operating at 160 or 240 MHz and performing at up to 600 DMIPS; in the MCU is also available an Ultra low power (ULP) co-processor. The MCU is equipped with 512kB of pSRAM (pseudo Static Random Access Memory) and 4MB of flash memory. The execution times have been measured through suitable software routines placed inside the classification code and they are not derived from the number of operations, time complexity or analytical analyses that are anyway present for the techniques' analysis in Subsection IV-C.2. Regarding the acquisition phase, the minimum required time to perform a single measurement on the 3 sensors has resulted in 372 ms. Whatever classification technique is adopted, the data acquisition time interval is the same (after that the CNN input buffer has been filled). Once the buffer has reached the steady state, a classification output is provided for each time–step as for the MLP. For this reason, an initial latency is needed for the CNN usage to acquire the w time–steps. Regarding the classification phase, 854 ms and 1 ms have been the elapsed times for CNN and MLP, respectively. In such a way, the CNN classification time results have an impact of 69% on the total elapsed time (acquisition + classification) while a impact lower than 1% is obtained for MLP. As regards the system response time, in terms of correct classification output, the system takes 2 minutes in average after the substance submission. Considering the distribution of correct classification during the whole submission, it has resulted that most of errors are concentrated in the initial phase. Finally, an evaluation of the energy consumption in a duty cycle has been carried out. The same rate used for dataset acquisition (one output every 2 seconds) has been maintained for the complete sequence. In such a way, after the acquisition and classification phases, the ESP32 turns in light–sleep mode, with a significant decrease of needed power. In Table 8 a summary of the required times and energy for each phase is reported. As result, an estimation of the energy consumption for each acquisition/classification cycle has been performed, obtaining 22.01 μWh and 73.04 μWh for MLP and CNN, respectively. Such values are particularly meaningful in order to obtain an estimation of the proposed system average lifetime in real battery operated scenarios.

TABLE 8. Elapsed time and energy Consumption by MCU during three different phases: DAcq = Data Acquisition; Infe = inference phase; Sleep = sleep phase between different acquisition. Values are reported for both MLP and CNN.

Phases	Power [mW]	*MLP* Elapsed [s]	*MLP* Energy [μWh]	*CNN* Elapsed [s]	*CNN* Energy [μWh]
DAcq	201.0	0.372	20.77	0.372	20.77
Infe	218.0	0.001	0.06	0.854	51.71
Sleep	2.6	1.627	1.18	0.774	0.56
Duty Cycle	/	2.000	22.01	2.000	73.04

V. CONCLUSIONS

A low–cost and low–power integrated system for pervasive indoor air monitoring is developed to detect the presence of contaminants. To this aim, both sensing, performed with a sensor array, and processing tasks are addressed in this work. In terms of processing, three different machine–learning based classification techniques have been tested. Due to obtained results and expected computational burden, MLP and CNN have been further investigated in terms of: i) computational complexity analysis, ii) ESP32 MCU implementation; iii) execution time and power consumption analyses. Six classes have been considered for the data acquisition and ML technique training. Four of them are reliably recognized, while acetone and alcohol detection is generally confused by the system. We explained the motivation of such confusion by noting chemical composition similarities between them. A possible solution to solve the problems is to employ other sensor typologies that are able to discriminate between them. The proposed solution is now under improvement and final deployment, to be released in the next future as rapid, flexible, distributed and reliable system to be used in indoor environments, especially in public or industrial buildings.

ACKNOWLEDGMENT

The authors are solely responsible for it and it does not represent the opinion of the Community and the Community is not responsible for any use that might be made of information contained therein. The authors gratefully acknowledge Sensichips s.r.l. for the support during the experimental phases and NVIDIA Corporation for the donation of the Titan Xp GPUs. Marco Ferdinandi contributed to the article until October 31, 2019.

REFERENCES

[1] M. Bentayeb, V. Wagner, M. Stempfelet, M. Zins, M. Goldberg, M. Pascal, S. Larrieu, P. Beaudeau, S. Cassadou, D. Eilstein, L. Filleul, A. Le Tertre, S. Medina, L. Pascal, H. Prouvost, P. Quénel, A. Zeghnoun, and A. Lefranc, "Association between long-term exposure to air pollution and mortality in france: A 25-year follow-up study," *Environ. Int.*, vol. 85, pp. 5–14, Dec. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0160412015300349

[2] M. Kampa and E. Castanas, "Human health effects of air pollution," *Environ. Pollut.*, vol. 151, no. 2, pp. 362–367, Jan. 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0269749107002849

[3] C. Santos, J. A. Jimenez, and F. Espinosa, "Effect of event-based sensing on IoT node power Efficiency. Case study: Air quality monitoring in smart cities," *IEEE Access*, vol. 7, pp. 132577–132586, 2019.

[4] L. Morawska *et al.*, "Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?" *Environ. Int.*, vol. 116, pp. 286–299, Jul. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0160412018302460

[5] N. Castell, F. R. Dauge, P. Schneider, M. Vogt, U. Lerner, B. Fishbain, D. Broday, and A. Bartonova, "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?" *Environ. Int.*, vol. 99, pp. 293–302, Feb. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0160412016309989

[6] Y. Li, W. Pang, C. Sun, Q. Zhou, Z. Lin, Y. Chang, Q. Li, M. Zhang, and X. Duan, "Smartphone-enabled aerosol particle analysis device," *IEEE Access*, vol. 7, pp. 101117–101124, 2019.

[7] P. Bruschi, G. Cerro, L. Colace, A. De Iacovo, S. Del Cesta, M. Ferdinandi, L. Ferrigno, M. Molinara, A. Ria, R. Simmarano, F. Tortorella, and C. Venettacci, "A novel integrated smart system for indoor air monitoring and gas recognition," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Jun. 2018, pp. 470–475.

[8] M. Ferdinandi, M. Molinara, G. Cerro, L. Ferrigno, C. Marroco, A. Bria, P. Di Meo, C. Bourelly, and R. Simmarano, "A novel smart system for contaminants detection and recognition in water," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Jun. 2019, pp. 186–191.

[9] G. Betta, G. Cerro, M. Ferdinandi, L. Ferrigno, and M. Molinara, "Contaminants detection and classification through a customized IoT-based platform: A case study," *IEEE Instrum. Meas. Mag.*, vol. 22, no. 6, pp. 35–44, Dec. 2019.

[10] Y. Deng, Z. Chen, X. Yao, S. Hassan, and J. Wu, "Task scheduling for smart city applications based on multi-server mobile edge computing," *IEEE Access*, vol. 7, pp. 14410–14421, 2019.

[11] X. Zhai, A. A. S. Ali, A. Amira, and F. Bensaali, "MLP neural network based gas classification system on zynq SoC," *IEEE Access*, vol. 4, pp. 8138–8146, 2016.

[12] Y. Sun, F. P.-W. Lo, and B. Lo, "A deep learning approach on gender and age recognition using a single inertial sensor," in *Proc. IEEE 16th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, May 2019, pp. 1–4.

[13] F. Röck, N. Barsan, and U. Weimar, "Electronic nose: Current status and future trends," *Chem. Rev.*, vol. 108, no. 2, pp. 705–725, Feb. 2008.

[14] S. Marco and A. Gutierrez-Galvez, "Signal and data processing for machine olfaction and chemical sensing: A review," *IEEE Sensors J.*, vol. 12, no. 11, pp. 3189–3214, Nov. 2012.

[15] G. Kesavaraj and S. Sukumaran, "A study on classification techniques in data mining," in *Proc. 4th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2013, pp. 1–7.

[16] G. Betta, C. Liguori, and A. Pietrosanto, "Uncertainty evaluation in algorithms with conditional statement," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 4, pp. 969–976, Aug. 2004.

[17] E. Esposito, S. De Vito, M. Salvato, V. Bright, R. L. Jones, and O. Popoola, "Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems," *Sens. Actuators B, Chem.*, vol. 231, pp. 701–713, Aug. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092540051630332X

[18] B. Mondal, M. S. Meetei, J. Das, C. Roy Chaudhuri, and H. Saha, "Quantitative recognition of flammable and toxic gases with artificial neural network using metal oxide gas sensors in embedded platform," *Eng. Sci. Technol., Int. J.*, vol. 18, no. 2, pp. 229–234, Jun. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2215098615000129

[19] S. De Vito, A. Castaldo, F. Loffredo, E. Massera, T. Polichetti, I. Nasti, P. Vacca, L. Quercia, and G. Di Francia, "Gas concentration estimation in ternary mixtures with room temperature operating sensor array using tapped delay architectures," *Sens. Actuators B, Chem.*, vol. 124, no. 2, pp. 309–316, Jun. 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925400506008562

[20] G. Huyberechts, P. Szecówka, J. Roggen, and B. W. Licznerski, "Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural network," *Sens. Actuators B, Chem.*, vol. 45, no. 2, pp. 123–130, Dec. 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925400597002839

[21] X. Li, L. Peng, Y. Hu, J. Shao, and T. Chi, "Deep learning architecture for air quality predictions," *Environ. Sci. Pollut. Res.*, vol. 23, no. 22, pp. 22408–22417, Nov. 2016.

[22] H. Karimian, Q. Li, C. Wu, Y. Qi, Y. Mo, G. Chen, X. Zhang, and S. Sachdeva, "Evaluation of different machine learning approaches to forecasting PM2.5 mass concentrations," *Aerosol Air Qual. Res.*, vol. 19, no. 6, pp. 1400–1410, 2019.

[23] P. Peng, X. Zhao, X. Pan, and W. Ye, "Gas classification using deep convolutional neural networks," *Sensors*, vol. 18, no. 2, p. 157, 2018. [Online]. Available: http://www.mdpi.com/1424-8220/18/1/157

[24] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," in *The Handbook Brain Theory Neural Networks*, M. A. Arbib, Ed. Cambridge, MA, USA: MIT Press, 1998, pp. 255–258. [Online]. Available: http://dl.acm.org/citation.cfm?id=303568.303704

[25] S. Hochreiter and J. Schmidhuber, "Long short–term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[26] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[27] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649.

[28] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0968090X15000935

[29] *Sensichips*. [Online]. Available: www.sensichips.com

[30] G. Cerro, M. Ferdinandi, L. Ferrigno, and M. Molinara, "Preliminary realization of a monitoring system of activated carbon filter RLI based on the SENSIPLUS microsensor platform," in *Proc. IEEE Int. Workshop Meas. Netw.*, Sep. 2017, pp. 1–5.

[31] G. Cerro, M. Ferdinandi, L. Ferrigno, M. Laracca, and M. Molinara, "Metrological characterization of a novel microsensor platform for activated carbon filters monitoring," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 10, pp. 2504–2515, Oct. 2018.

[32] *Humidity Modules Sensors*. [Online]. Available: https://www.ist-ag.com/en/products-services/humidity-modules-sensors

[33] F. Ricciardella, S. Vollebregt, T. Polichetti, M. Miscuglio, B. Alfano, M. L. Miglietta, E. Massera, G. Di Francia, and P. M. Sarro, "Effects of graphene defects on gas sensing properties towards NO_2 detection," *Nanoscale*, vol. 9, no. 18, pp. 6085–6093, 2017, doi: 10.1039/C7NR01120B.

[34] B. Alfano, T. Polichetti, M. Mauriello, M. L. Miglietta, F. Ricciardella, E. Massera, and G. Di Francia, "Modulating the sensing properties of graphene through an eco-friendly metal-decoration process," *Sens. Actuators B, Chem.*, vol. 222, pp. 1032–1042, Jan. 2016.