IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 4, 2018, accepted January 4, 2019, date of publication January 14, 2019, date of current version February 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892754

Mixture Gases Classification Based on Multi-Label One-Dimensional Deep Convolutional Neural Network

XIAOJIN ZHAO[1], (Member, IEEE), ZHIHUANG WEN[1], XIAOFANG PAN[2], (Member, IEEE), WENBIN YE[1], (Member, IEEE), AND AMINE BERMAK[3,4], (Fellow, IEEE)

[1]College of Electronic Science and Technology, Shenzhen University, Shenzhen 518060, China
[2]College of Information Engineering, Shenzhen University, Shenzhen 518060, China
[3]Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong
[4]College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar.

Corresponding author: Xiaofang Pan (eexpan@szu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61504087, in part by the Kongque Technology Innovation Foundation of Shenzhen under Grant KQJSCX20170727101037551, and in part by the Fundamental Research Foundation of Shenzhen under Grant JCYJ20170818101906654 and Grant JCYJ20170302151209762.

ABSTRACT In this paper, we present a novel one-dimensional deep convolutional neural network (1D-DCNN) with a multi-label-way-based algorithm for comprehensively and automatically extracting features and classifying mixture gases. Although a number of pattern recognition methods have been used to analyze the mixed gases, the performances of these methods highly depend on the hand-crafted feature engineering. By contrast, the proposed implementation, based on one-dimensional convolution, is capable of automatically extracting features and distinguishing the individual component of binary mixture gases composed of ethylene, CO, and methane. To the best of our knowledge, the proposed 1D-DCNN algorithm is first applied in the mixture gases' recognition. In addition, the proposed 1D-DCNN with multi-label way not only significantly reduces the label dimension but also quantifies the probability of each component in mixed gases. Compared with the conventional pattern recognition algorithms including support vector machine, artificial neural network, k-nearest neighbor, and random forest, the proposed 1D-DCNN exhibits a higher recognition accuracy (96.30%) based on our extensive experimental results using ten-fold cross validation.

INDEX TERMS Mixture gases recognition, deep convolutional neural network, multi-label classification.

I. INTRODUCTION

In the past few years, the recognition of mixture gases has attracted extensive research interests in a wide range of industrial applications, including health care, environment monitoring, food freshness control, to name a few [1]–[4]. It is known that the hazardous gases in industrial applications are usually not pure but in the form of mixed gases. Therefore, in order to reduce the danger of the hazardous mixture gases, the accurate recognition of different compositions in the mixture gases is of great importance. As far as we know, electronic nose is widely adopted as an effective way to realize the recognition of mixture gases, which well-mimics the mammals' olfactory system [5], [6]. There are mainly three building blocks in the modern electronic nose: gas sensing array, pre-processing module and signal processing unit. For a given gas sensing array, the signal processing unit always plays a key role for recognizing different gas compositions with high accuracy.

Recent years have witnessed many efforts devoted to further improve the recognition accuracy of the mixture gases. Sunny *et al* demonstrated a recognition system for binary mixture gases of acetone and 2-propanol. The principal component analysis (PCA) is first used to extract the mixture gases' steady response features then the artificial neural network (ANN) and support vector machine (SVM) are adopted for gas classification [7]. In [8], a hand-crafted feature of normalized difference sensor response transformation (NDSRT) is reported to convert the raw signature responses to virtual multi-sensor response, which further improves the accuracy of gas classification. In [9], different pattern recognition methods including ANN, SVM and extreme learning machine (ELM) are exploited to recognize

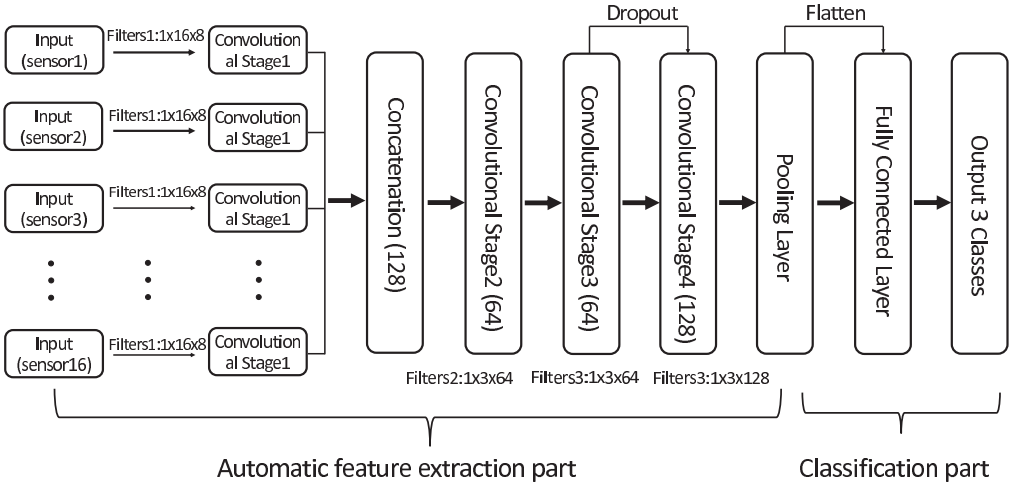


FIGURE 1. Overall architecture of the proposed 1D-DCNN algorithm.

formaldehyde with the background of ethanol, acetone and toluene in air, where their input are the data from the first two principal components of raw responses obtained from PCA. Moreover, Yan *et al* applied support vector machine recursive feature elimination (SVM-RFE) to select effective transient features extracted from gas sensor responses in terms of magnitude, difference, derivative, second order derivative, integral, etc [10]. In [12], the responses of binary gas mixtures of acetone, ethanol and trichloroethylene, are first preprocessed by a sliding window algorithm to obtain valid data set, then directly classified by ANN without any feature extraction. However, the reported accuracy is limited to be 84.5%. From the above literatures, there is one common drawback that these traditional gas recognition methods' performances highly depend on the quality of the hand-crafted features. As the mixture gases' raw response is not linear with several gases, different feature extraction techniques including PCA and statistical feature method (e.g. standard deviation, mean and variance) are always necessary to generate the hand-crafted features [10], [11]. Meanwhile, the hand-crafted features are various and not applicable to the wide range of classification tasks. Therefore, it is quite challenging to propose a uniform effective feature extraction method for the hand-crafted features.

In this paper, we propose a one-dimensional deep convolutional neural network (1D-DCNN) based algorithm dedicated to the classification of mixture gases. The proposed 1D-DCNN algorithm, which is an end-to-end model that automatically extract feature without human interventions, is capable of comprehensively extracting and classifying different features of target gases directly from the raw data. This is quite different from the traditional gas recognition algorithms, where massive priori knowledge of the mixture gases' features and human effort in feature extraction are always inevitable. In addition, the proposed algorithm includes four convolutional stages, a dropout layer, a pooling layer and two fully-connected layers. Meanwhile, each convolutional stage consists of three layers: one 1D convolutional layer, one batch normalization layer and one rectified linear unit (ReLU) layer, which are utilized to automatically extract features from the raw data. The main contributions of this work are summarized as follows: 1) as far as we know, the proposed 1D-DCNN is firstly applied in the mixture gases recognition; 2) in contrast to most previous mixture gases classification algorithms that process the mixed gases as a new gas, the proposed 1D-DCNN adopts multi-label as the final output, which not only significantly decreases the label dimension but also presents the probability of each gas component intuitively; 3) as the different gas components are independent existence, the 1D-DCNN's final loss is the mean of mixed components' log loss which can be modified by the binary log loss. The rest of this paper is organized as follows. In Section II, the architecture of the proposed algorithm is illustrated and detailed. Experimental results and discussions are presented in Section III. Finally, the concluding remarks are drawn in Section IV.

II. PROPOSED ALGORITHM FOR MIXTURE GASES RECOGNITION

In this section, the proposed deep learning framework based 1D-DCNN algorithm is elaborated. As shown in Fig. 1, the proposed implementation's overall architecture mainly contains two parts: feature extraction part and classification part. In the first part, we combine convolutional operation, batch normalization operation and activation operation together as the convolutional stage, with different parameters used at different stages. And the other part is the multi-layer perception (MLP) part.

A. FEATURE EXTRACTION AND CLASSIFICATION

Feature extraction - In contrast to traditional CNN, the gas sensors' responses for mixture gases are multiple 1D subsequences rather than 2D image pixels. Compared with the image recognition, the responses of gas sensor array are

horizontally relevant instead of both vertically and horizontally correlated in image sensing array [13], [14]. Hence, 1D convolutional operation is used on each sensor response. Then the learned features are concatenated together to further extract the lumped feature of different sensors, which can dramatically improve the selectivity and sensitivity of the sensor array. Successively, we adopt the relatively small filters (64 or 128) and increase the model layers to promote the generalization of model [15]. Additionally, between the third and fourth stage of the convolution, a dropout layer is inserted to decrease the operating neural units of the third convolutional stage, which can avoid overfitting and improve computation efficiency during the training phase. In the whole framework, we only employ one pooling layer (average pooling) that down-samples a number of extracted features by one third to remove redundant features acquired from the fourth convolutional stage. Since the gas sensors' responses are continuous and relevant in time domain, the average pooling layer is used to process the values at different neighboring positions of the feature maps and guarantee the continuity of the down-sampled features.

Classification - The rest layers in 1D-DCNN are fully connected to form MLP, which is used for classifying the extracted features. As the computational complexity of MLP parameters dominates the whole 1D-DCNN computation, in order to keep the computational complexity at a suitable level, the numbers of fully-connected layers' neural units are reduced to 64 and 32. In addition, ReLU functions are employed as their activation functions and a dropout operation is inserted before the first fully-connected layer to improve the computation efficiency and avoid overfitting. Finally, in the output layer, we apply 3 neurons and sigmoid function to output 3 multi-label class probabilities (Ethylene, CO, Methane), which is elaborated in the following subsections.

B. CONVOLUTIONAL STAGE

The convolutional stage of the proposed 1D-DCNN algorithm comprise a few operations, such as convolutional operation, batch normalization (BN) operation and activation operation. As described in previous subsection, convolutional operations are a series of one-dimensional filters. The sizes of four convolutional kernels are $1 \times 16 \times 8$, $1 \times 3 \times 64$, $1 \times 3 \times 64$ and $1 \times 3 \times 128$, respectively, where 8, 64, 64 and 128 are the numbers of kernels in four convolutional stages. The detailed 1D convolutional operation is illustrated in Fig. 2. Similar to the traditional 2D convolution, we assume that input series A has dimension of $(1, N_a)$ and kernel B has dimension of $(1, N_b)$, the equation of the 1D convolution can be expressed as follows:

$$C(j) = \sum_{n=1}^{N_b} A(j+n) \times B(n) \tag{1}$$

where $1 \leq j \leq N_a - N_b + 1$ and $C(j)$ represents the j-th convolutional result. As the distribution of the initialized

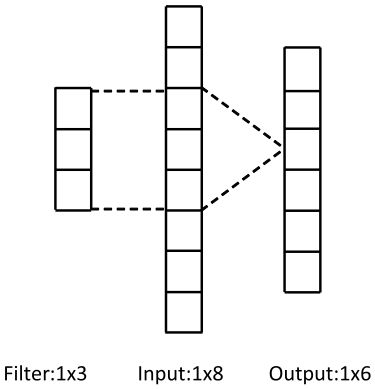


FIGURE 2. Detailed 1D convolutional operation.

weights strongly influences the model performance and its generalization ability, xavier initialization is utilized as the weights initialization method [16], which is an effective neural network initialization method and is capable of making the variance of every layer's output equal in order to obtain the optimized information flow. The uniformly distributed xavier initialization can be expressed as:

$$W \sim U\left[-\sqrt{\frac{6}{m+n}}, \sqrt{\frac{6}{m+n}}\right] \tag{2}$$

where m and n are the numbers of input units and output units in each layer.

BN and activation operation - Batch normalization can significantly improve the performance of the proposed layer with nonlinear sigmoid function [17]. Consequently, BN layer is inserted before activation in every convolutional stage. It is worthy to mention that adding BN operation not only accelerates the training process, but also avoids overfitting especially for relatively small dataset. Similar to the traditional CNN, here ReLU function is adopted as our activation function in convolutional stage.

C. MULTI-LABEL CLASSIFICATION

With the increment of the mixture gases' components, the gas combinations increase rapidly as well and the label dimension becomes large if we use one-hot representation. For example, provided with the mixed gases having n gas components, the label dimension is equal to 2^n using one-hot representation; while its dimension is equal to n using a multi-label way. As a result, multi-label classification is adopted to reduce the overall computation amount caused by the large label dimension. There are two mainstream methods for multi-label classification: a) problem transformation method, and b) algorithm adaptation method [18]. In this paper, we adopt problem transformation method to divide the multi-label classification problem into several binary classifier problems. As shown in Table 1, the dataset used in Section III has 6 multi-label combinations, where '1' means the corresponding gas exists and [0, 0, 0] means there is no target gas and only the reference gas (i.e. air) exists. Here we divide this multi-label classification problem into three binary

TABLE 1. Example of multi-label dataset.

No.	Ethylene	CO	Methane
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1
5	1	1	0
6	1	0	1

classifier problems with single-label: {Ethylene, ¬Ethylene}, {CO, ¬CO} and {Methane, ¬Methane} (¬ represents the label for negative examples). Since the labels of the mixture gas datasets are independent, the transformation of the three binary classifiers can be implemented by sigmoid function, which is easier to be combined with CNN-based method.

Additionally, the output of 1D-DCNN is described in Eq. (3), where y_{pre} represents the prediction of target gases, W^T represents the second hidden layer's weight. The output of sigmoid function is employed as the result of these three classifiers and a threshold value of 0.5 is set to produce the predicting labels, where the output probability over 0.5 is set to 1 and lower than 0.5 is set to 0. Moreover, regarding the loss value of this model, inspired by [18], we modify the binary log loss of three classifiers as Eq. (4), where $y(c_i)$ is the true label, N is the classifier number, $y_{pre}(c_i)$ and $q(c_i)$ are the relative and negative probabilities of the class (c_i) in the ith classifier, respectively. Finally, we average the log loss values in these binary classifiers as the final log loss value.

$$y_{pre} = sigmoid(W^T h + b) \tag{3}$$

$$Log\ loss = -\frac{1}{N}\sum_{i=1}^{N}\left[y(c_i)\log y_{pre}(c_i) + (1 - y(c_i))\log q(c_i)\right] \tag{4}$$

III. EXPERIMENTAL RESULTS AND COMPARISON

In order to verify the effectiveness of this model for mixture gas classification, extensive experiments were conducted with the public available dataset [19]. This dataset contains two binary mixture gases: Ethylene and Methane in air, Ethylene and CO in air. The response signal data was acquired from 16 metal oxide (MOX) sensors, which are exposed to various gas conditions and consist of four different types (TGS-2600, TGS-2602, TGS-2610, TGS-2620). The operating voltage was set to be 5V and the response signals of sensors were acquired at a sampling frequency of 100Hz continuously for 12 hours. As the raw dataset is labeled in its concentration types with $4178504 + 4208261 = 8386765$ instances, where 4178504 represents the instances between Ethylene and Methane and 4208261 represents the instances between Ethylene and CO, we manually label the category in manner of multi label and split the raw dataset into 593 samples that represent the same class respectively. Furthermore, we sampled the dataset down to 10Hz and randomly split the whole samples into 80% training and 20% testing sets (training sets: $593 \times 0.8 \approx 475$, testing sets: $593 \times 0.2 \approx 118$). Due to the unavoidable drifts of MOX sensors array, the baselines of mixed gases' reaction and recovery might deviate. Therefore, we selected response signals in 10s (100 points) with the biggest gradient in every sample as the model input. The gradient of the 10s response can be calculated as follows:

$$\nabla G = \frac{|G(t_i) - G(t_{i+10})|}{G(t_i)} \tag{5}$$

where t_i is the time point in every sample. In Fig. 3, for the Ethylene sample with the reaction phase between 500s and 600s, we choose the duration between 526s and 536s as the model input.

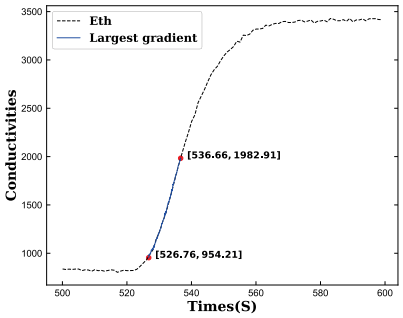


FIGURE 3. The reaction with Ethylene between 500s and 600s.

A. FIGURES OF MERIT FOR PERFORMANCE EVALUATION

As described in previous section, this multi-label mixture gases classification task is converted to several binary classification tasks, which belong to the label-based measurement. To evaluate the performance of the classifier in mixture gases classification, three most important figures of merit, including accuracy, log loss and confusion matrix are detailed as follows [18], [20]:

1) For test set evaluation, accuracy is the first figure of merit to evaluate the performance. In Eq. (6), D is the test data set, consisting of $|D|$ multi-label examples (x_i, Y_i) (i=1 . . . $|D|$, $Y_i \subseteq \{Air, Eth, CO, Met, Eth-CO, Eth - Met\}$, where $Eth - CO$ represents the label of the Ethylene/CO mixture and $Eth - Met$ represents the Ethylene/Methane mixture), f is a classifier and $P_i = f(x_i)$ is the predicted label set by the classifier f. The model will be better if the test set accuracy is larger.

$$Accuracy = \frac{1}{|D|}\sum_{i=1}^{|D|}\frac{|Y_i \cap P_i|}{|Y_i \cup P_i|} \tag{6}$$

2) Log loss is described in Eq. (3), which is used in logistic regression that exists in 1D-DCNN output layer. Besides, log loss values represent the training error and test error of the train set and test set, respectively. If the log loss value is smaller, the model fits the data set better. Furthermore, if the test log loss is very close to the training log loss, the over-fitting or under-fitting phenomenon can be well-avoided.
3) In the field of machine learning (especially pattern recognition), the confusion matrix is a specific matrix

used to present visualization of an algorithm's performance. Compared with the other figures of merit (accuracy and F1 score), the confusion matrix enables more detailed analysis of the correct classification and the mislead results for a classifier. Moreover, the confusion matrix is also known as the error matrix, where each column of the matrix represents the instance in a predicted class while each row represents the true instance of the corresponding class. In this paper, we adopt a normalized confusion matrix as the third figure of merit. With the diagonal value in each row closer to 1, more accurately predicted classes can be produced.

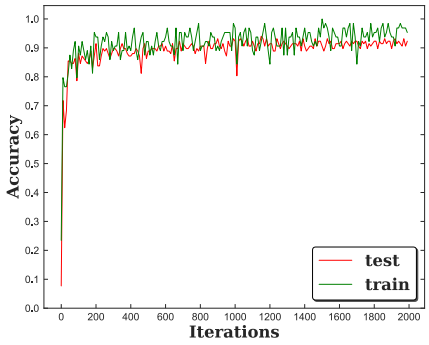
B. EXPERIMENTAL RESULTS

Here the proposed 1D-DCNN algorithm is implemented with *TensorFlow*, the detailed steps are summarized as follows:

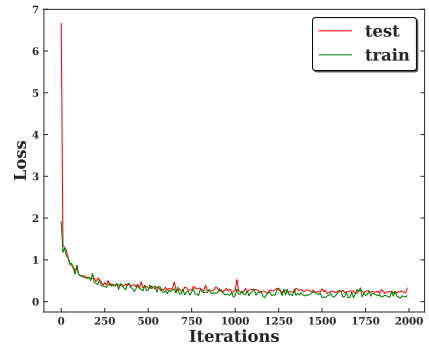
1) Optimization: we apply mini-batch stochastic gradient descent with Adam optimizer [17]. The batch size is set to 64. As described in previous section, Batch Normalization operations are adopted in feature extraction part and classification part.
2) Regularization: the dropout operations with keep-probability equal to 0.25 are adopted in convolutional stage 3 and the second hidden layer of classification.
3) Hyperparameters: In classification part, we apply two hidden layers with 128 and 64 neurons, respectively. For learning rate, we first manually tune the fix learning rate between 0.00001 and 0.001 while it is difficult for model convergence and stability. Finally, we apply the exponential-decay learning rate, where its initial learning rate is 0.01 and its decay step is 10.

The performance of 1D-DCNN is evaluated by log loss and accuracy in train set and test set, while evaluated by confusion matrix only for test set. As described in previous section, with larger accuracy and smaller log loss, the better performance of model can be achieved. During 2000 iterations with stochastic mini-batches, the test accuracy and train accuracy are almost equal after 1000 iterations and the highest test accuracy is reported to be 95.73%, as depicted in Fig. 4 (a). While in Fig. 4 (b), the log loss in testing set and training set descend to less than 0.5 after 500 iterations and finally stabilize at around 0.2. These results suggest that the underfitting or overfitting phenomenon is well-avoided using this proposed model. Finally, the normalized confusion matrix of 1D-DCNN is indicated in Fig. 5 (a), where the predicted and true classes (Air, Eth, CO, Met, Eth-CO, Eth-Met) are in column and row of each table respectively, the diagonal values represent the normalized numbers of predicted classes equal to true classes and the corrective rates of the two mixture gases Eth-CO and Eth-Met are close to 0.90. The corrective rate of each predicted class is fully visualized by the confusion matrix.

In addition, K-fold cross validation is a universal means for evaluating the results acquired from pattern recognition methods. Here we use 10-fold cross validation (CV) to evaluate the proposed 1D-DCNN and the previously reported methods,



(a)



(b)

FIGURE 4. (a) the curves of train and test accuracy; (b) the curves of train and test loss.

where the whole set is randomized first then divided into ten segments with equal size. Moreover, with each segment used as the testing dataset and the other nine segments combined as the training set, we can have altogether 10 different training sets. Finally, the 10 different sets are fed into the proposed method successively to generate 10 maximized accuracies, which are averaged to yield the final recognition accuracy of 96.30%.

C. COMPARISON

For comparison, we considered four previous approaches: SVM, ANN, KNN and RF. Concretely, in the past few years, a number of approaches have been proposed for binary mixture gas recognition. In [7], SVM is applied for classifying two volatile organic compounds (VOC) gases. Before neural networks, SVM has been the mainstream solution for pattern recognition. As a result, we use SVM as our first candidate for comparison. Moreover, ANN is widely-adopted for recognizing the mixture gases as well [9], [21]. Hence, ANN is used as another method for comparison, where the neuron number of its two hidden layers are 300 and 100, respectively. Additionally, the comparison is further extended to the general machine learning method KNN [22] and ensemble learning method RF [23], which are quite popular for recognizing pure gas. As the inputs for these previous methods are usually one-dimensional vectors, we unfolded the raw input data into one-dimensional vector (16×100) and one-hot

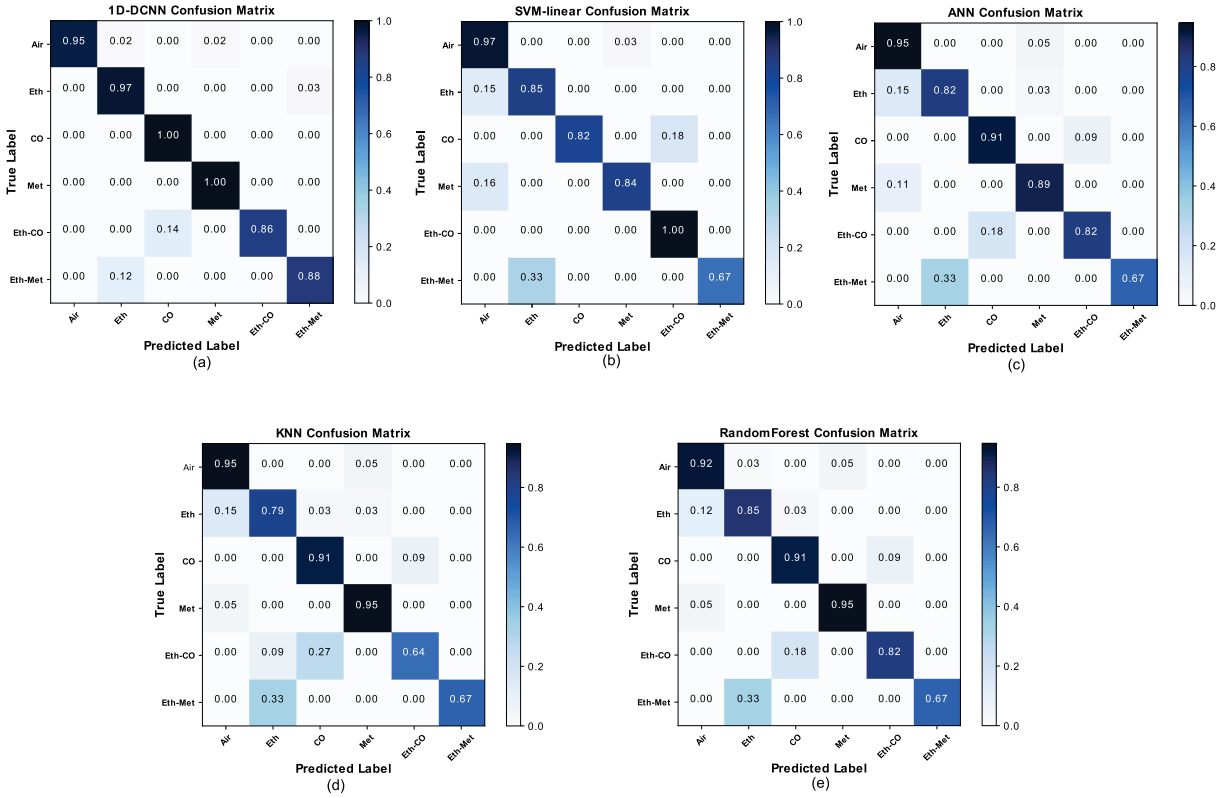


FIGURE 5. Normalized confusion matrix comparison, where its column instances are predicted classes and row instances are true classes: (a) the normalized confusion matrix of 1D-DCNN; (b) the normalized confusion matrix of SVM with linear kernel; (c) the normalized confusion matrix of ANN; (d) the normalized confusion matrix of KNN; (e) the normalized confusion matrix of RF.

TABLE 2. The performance comparison among different principal components processed by PCA.

Algorithms	20 components	50 components	100 components	200 components	300 components
SVM-linear	86.28%	87.44%	**87.45%**	87.29%	87.29%
SVM-rbf	82.93%	82.93%	82.93%	82.93%	82.93%
SVM-poly	82.23%	86.16%	**88.14%**	86.08%	85.56%
RF	88.18%	**88.69%**	88.02%	86.53%	84.51%
KNN	80.28%	80.45%	80.45%	80.45%	80.45%
ANN	83.84%	84.36%	82.51%	**85.85%**	82.50%

representation is adopted for labeling mixture gases in these previous methods. Meanwhile, the input vectors are the first n-components (n is the dimension of components) preprocessed with principal component analysis (PCA) to provide features for the method to be evaluated. Finally, 10-fold cross validation is also adopted for these previous methods as the proposed 1D-DCNN.

CV Comparison: As shown in Fig. 6, the preprocessed data is extracted using PCA, where $X1$, $X2$ and $X3$ represent the eigenvectors in PCA. The first two principal components and the first three components are plotted with two-dimensional scatter and three-dimensional scatter, respectively. Different color scatters represent different gases decomposed by PCA. In Fig. 6 (b), the scatters of Eth, CO and Eth-CO cluster together are difficult to be classified by these previous methods. Therefore, higher dimensional decomposition can increase the eigenvectors' representativeness and is necessary to the following classification. As shown in Table 2, we compare the four previous algorithms' performances under different principal components. Concretely, SVM with linear kernel (SVM-linear) and SVM with poly kernel (SVM-poly) can achieve the best performance with

TABLE 3. The comparison of 10-fold cross-validation recognition accuracy between 1D-DCNN and the other four previous methods.

Algorithms	SVM	ANN	KNN	RF	**1D-DCNN**
Accuracies	87.45%	85.85%	80.45%	88.69%	**96.30%**

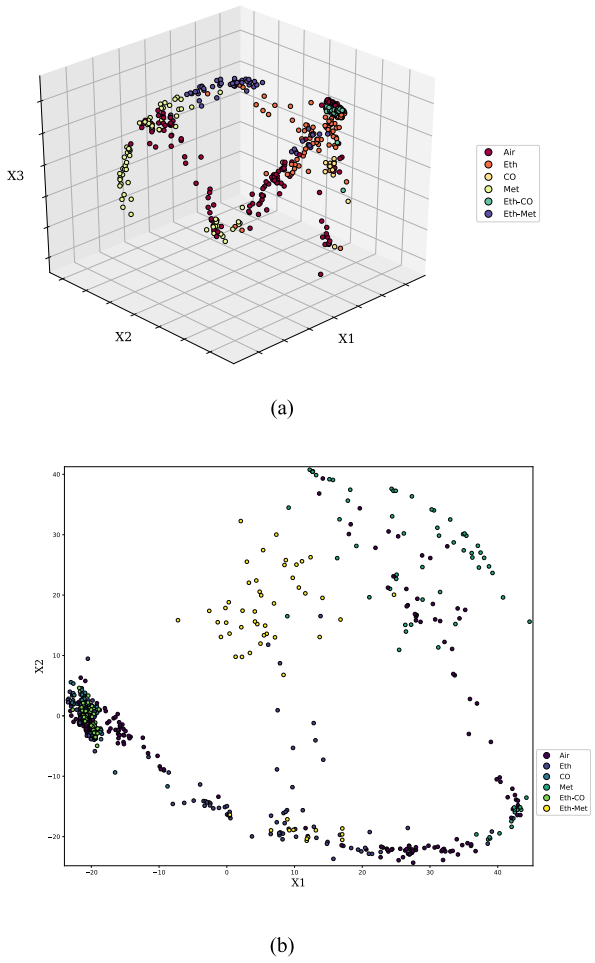


FIGURE 6. (a) the first three principal components; (b) the first two principal components.

100 principal components and RF can achieve the best performance with 50 principal components. In addition, ANN algorithm requires 200 dimensional eigenvectors to produce the best performance, while the SVM with radial basis function (SVM-rbf) and KNN are insensitive to the number of principal components. For our proposed 1D-DCNN algorithm, as described in previous section, it can automatically extract the effective features by a serious of convolutional computations, which removes the need of hand-crafted features generated by PCA.

Moreover, Table 3 summarizes the comparison of CV recognition accuracies between the 1D-DCNN and the four traditional methods, where 100, 200, 50 and 20 principal components are selected as the input vectors of SVM-linear, ANN, RF and KNN, respectively. It is indicated that the KNN's recognition accuracy is the lowest, the result of ANN with two hidden layers is 85.85% and the recognition accuracies of RF and SVM are quite close to each other. Meanwhile, the recognition accuracy of the proposed 1D-DCNN largely outperforms the aforesaid four algorithms, which can be attributed to the meaningful and representative features generated by the deep network layers and lots of neurons.

Confusion Matrix Comparison: in Fig. 5, five confusion matrices illustrate the performance comparison of 1D-DCNN and the other four previous algorithms. It is observed that the four previous algorithms present favorable corrective rates in each pure gas class while 1D-DCNN shows superior performance. However, the previous algorithms cannot exhibit excellent performance for two mixture gas classes especially the mixture gases between Ethylene and Methane. On the contrary, the corrective rate of 1D-DCNN can still achieve 0.90. As a result, the proposed 1D-DCNN method outperforms all the previous algorithms for each pure gas class as well as the mixture gas classes.

IV. CONCLUSION

In this paper, we present a novel multi-label 1D-DCNN-based algorithm for classifying binary mixture gases among Ethylene, CO and Methane. To the best of our knowledge, the proposed 1D-DCNN is first exploited for mixture gases classification, which automatically classifies the features directly extracted from the raw data set. Meanwhile, the proposed 1D-DCNN is an end-to-end classifier, which is capable of significantly simplifying the extraction process of more meaningful and high level features by the trainable neural networks. In addition, compared with all previously demonstrated algorithms, the 1D-DCNN processes the raw valid data set in a multi-label way, which greatly reduces the dimension of the label space. According to the extensive experimental results, it is indicated that the corrective rates of the proposed 1D-DCNN in each pure gas as well as mixture gases outperform the other algorithms. Moreover, the comparison is further extended to 10-fold cross validation. It is observed that the average recognition accuracy of the proposed 1D-DCNN significantly outperforms the conventional methods of SVM, MLP, KNN and RF. As a result, the proposed algorithm is suitable for a wide range of classification applications of the mixture gases, where it is always quite challenging to obtain effective artificial feature.

REFERENCES

[1] A. Somov, A. Karelin, A. Baranov, and S. Mironov, "Estimation of a gas mixture explosion risk by measuring the oxidation heat within a catalytic sensor," *IEEE Trans. Ind. Electron.*, vol. 64, no. 12, pp. 9691–9698, Dec. 2017.

[2] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Application of electronic nose systems for assessing quality of medicinal and aromatic plant products: A review," *J. Appl. Res. Med. Aromatic Plants*, vol. 3, no. 1, pp. 1–9, 2016.

[3] A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, and J. B. B. Rayappan, "Electronic noses for food quality: A review," *J. Food Eng.*, vol. 144, pp. 103–111, Jan. 2015.

[4] A. C. Romain and J. Nicolas, "Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview," *Sens. Actuators B, Chem.*, vol. 146, no. 2, pp. 502–506, Apr. 2010.

[5] S. Ampuero and J. O. Bosset, "The electronic nose applied to dairy products: A review," *Sens. Actuators B, Chem.*, vol. 94, pp. 1–12, Aug. 2003.

[6] S.-W. Chiu and K.-T. Tang, "Towards a chemiresistive sensor-integrated electronic nose: A review," *Sensors*, vol. 13, no. 10, pp. 14214–14247, 2013.

[7] Sunny, V. Kumar, V. N. Mishra, R. Dwivedi, and R. R. Das, "Classification and quantification of binary mixtures of gases/odors using thick-film gas sensor array responses," *IEEE Sensors J.*, vol. 15, no. 2, pp. 1252–1260, Feb. 2015.

[8] A. Mishra, N. S. Rajput, and G. Han, "NDSRT: An efficient virtual multi-sensor response transformation for classification of gases/odors," *IEEE Sensors J.*, vol. 17, no. 11, pp. 3416–3421, Jun. 2017.

[9] L. Zhao, X. Li, J. Wang, P. Yao, and S. A. Akbar, "Detection of formaldehyde in mixed VOCs gases using sensor array with neural networks," *IEEE Sensors J.*, vol. 16, no. 15, pp. 6081–6086, Aug. 2016.

[10] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sens. Actuators B, Chem.*, vol. 212, pp. 353–363, Jun. 2015.

[11] M. F. Adak, M. Akpinar, and N. Yumusak, "Determination of the gas density in binary gas mixtures using multivariate data analysis," *IEEE Sensors J.*, vol. 17, no. 11, pp. 3288–3297, Jun. 2017.

[12] A. Özmen, F. Tekce, M. A. Ebeoğlu, C. Taşaltin, and Z. Z. Öztürk, "Finding the composition of gas mixtures by a phthalocyanine-coated QCM sensor array and an artificial neural network," *Sens. Actuators B, Chem.*, vol. 115, no. 1, pp. 450–454, 2006.

[13] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *J. Syst. Eng. Electron.*, vol. 28, no. 1, pp. 162–169, Feb. 2017.

[14] Z. Cui, W. Chen, and Y. Chen. (2016). "Multi-scale convolutional neural networks for time series classification." [Online]. Available: https://arxiv.org/abs/1603.06995

[15] M. Nielsen. (2017). *Neural Networks and Deep Learning*. [Online]. Available: http://neuralnetworksanddeeplearning.com/

[16] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249–256, May 2010.

[17] I. Goodfellow, Y. Bengio, A. Courville, and F. Bach, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[18] G. Tsoumakas and I. Katakis, "Multi-label classification: An overview," *Int. J. Data Warehousing Mining*, vol. 3, no. 3, pp. 1–13, 2007.

[19] J. Fonollosa, S. Sheik, R. Huerta, and S. Marco, "Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring," *Sens. Actuators B, Chem.*, vol. 215, pp. 618–629, Aug. 2015.

[20] D. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.

[21] B. Mumyakmaz, A. Özmen, M. A. Ebeoğlu, and C. Taşaltin, "Predicting gas concentrations of ternary gas mixtures for a predefined 3D sample space," *Sens. Actuators B, Chem.*, vol. 128, no. 2, pp. 594–602, 2008.

[22] W. Li, H. Leung, C. Kwan, and B. R. Linnell, "E-nose vapor identification based on Dempster–Shafer fusion of multiple classifiers," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 10, pp. 2273–2282, Oct. 2008.

[23] Q. Li, Y. Gu, and N.-F. Wang, "Application of random forest classifier by means of a QCM-based E-Nose in the identification of Chinese liquor flavors," *IEEE Sensors J.*, vol. 17, no. 6, pp. 1788–1794, Mar. 2017.

XIAOJIN ZHAO (S'07–M'10) received the B.Sc. degree from the Department of Microelectronics, Peking University, in 2005, and the Ph.D. degree from the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology (HKUST), in 2010. After one-year Postdoctoral research work at HKUST, he joined the College of Electronic Science and Technology, Shenzhen University, where he is currently an Associate Professor. His research interests include CMOS monolithic polarization image sensor, gas sensor, and their related hardware security when applied to the field of the Internet of Things.

ZHIHUANG WEN received the B.Sc. degree from the College of Electronic Science and Technology, Shenzhen University, Shenzhen, China, in 2016, where he is currently pursuing the master's degree. His research interests include the CMOS analog circuit design and algorithm development for sensor applications.

XIAOFANG PAN (S'12–M'16) received the B.Eng. degree in electronic science and engineering from Southeast University, Nanjing, China, in 2010, and the Ph.D. degree from The Hong Kong University of Science and Technology, Hong Kong, in 2015. In 2015, she joined the College of Information Engineering, Shenzhen University, where she is currently an Assistant Professor. Her research interests include gas sensor/electronic nose fabrication, the related gas species classification, and on-chip system design.

WENBIN YE (S'12–M'14) received the B.Sc. degree in microelectronics from Sichuan University, Chengdu, China, in 2009, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2014. He is currently an Associate Professor with Shenzhen University. His research interests include digital filter design, nonuniform sampled data processing, implementation, and image processing.

AMINE BERMAK (F'13) received the master's and Ph.D. degrees in electrical and electronic engineering (microelectronics and microsystems) from Paul Sabatier University, Toulouse, France, in 1994 and 1998, respectively. In 2002, he joined the Electronic and Computer Engineering Department, The Hong Kong University of Science and Technology (HKUST), where he is currently a Professor. He is also the Founder and the Leader of the Smart Sensory Integrated Systems Research Laboratory, HKUST. His research interests include VLSI circuits and systems, packaging technologies, CMOS image sensors, smart vision systems, electronic nose and olfactory systems, and VLSI implementation of pattern recognition algorithms.