Chapter 2

Materials, Methods, and Technology Background

2.1 Materials

As discussed in Chapter 1, metal oxides-based semiconductors have been quite popular in the gas sensing community. Except for other existing gas sensors, metal-oxide (MOX)-based gas sensors exhibit gas sensing phenomena by changing the electrical properties of the sensing materials. Such gas sensors are cheaper and provide high selectivity. In our published works to date, we have used three publicly available datasets along with one dataset captured in our own department's laboratory. All these datasets had been recorded using the gas sensor array consisting of MOX-based gas sensors. Before discussing these datasets, complete detail of MOX gas sensors would be beneficial.

2.1.1 Metal Oxide Semiconductors

In comparison to elemental semiconductors, metal oxide semiconductors show unique properties due to the disparity of charge carrier transport caused by the interaction between the metal and oxide orbitals. In metal oxides, the electrons have smaller effective masses than holes, resulting in high mobility to improve carrier transport. Thus, such metal oxide semiconductors, e.g., SnO_2 and ZnO, represent n-type conductivity. However, p-type conductivity can also be achieved in metal oxides if the holes have an effective mass significantly smaller than electrons. As we found in the literature, since 1993, Nickel Oxide (NiO) has been known as the first p-type metal oxide semiconductor [145]. Besides NiO, Cu_2O and $CuMO_2(M = Al, Ga, orIn)$ are also p-type metal oxide semiconductors [146].

2.1.2 Working of MOX Gas Sensors

In MOX gas sensors, the sensing material is printed on the base material of the device. This arrangement is mounted over a heater used to induce the thermal excitation for sensing material. This heater generates a temperature of several hundred degrees of Celsius. Thus, being thermally excited, the covalent bonds are broken, resulting in free electrons in the sensing material. These free electrons flow within the sensing material and provide a measurable current. However, this phenomenon occurs when the sensing material is placed in an inert environment.

On the other hand, if the sensing device is placed in the clean air, the present oxygen is adsorbed on the sensing material. The adsorbed oxygen traps the free electrons of the sensing material, causing an increase in the resistance of the sensing material. This phenomenon tries to stop the flow of free electrons. When this device is exposed to the reducing gases/odors (e.g., methane, propane, etc.), it reacts with

the adsorbed oxygen on the sensing material. This reaction makes the free electrons trapped by the adsorbed oxygen accessible again within the sensing material. It reduces the resistance of sensing material, allowing more flow of electrons resulting in significant current. This increase in the flow of free electrons is directly proportional to the concentration of exposed gases/odors.

2.1.3 Responses of MOX Gas Sensors

As we have studied in the last subsection, the change in resistance of the sensing material is directly proportional to the change in concentration of the exposed gas/odor. To detect the gas, we observe this change in resistance. While detecting the exposed gas/odor, the observed change in resistance (decrease or increase), depends on the types of conductivity of sensing material (n-type or p-type) and the nature of exposed gas/odor (reducing or oxidizing). Although, in each case, the change in resistance is correlated with the concentration of gas/odor. A schematic response curve for the MOX gas sensor has been shown in Fig. 2.1. The MOX gas sensors detect the gases/odors by measuring the sensing materials' electrical properties (resistance, conductance, etc.).

The characterization of gas sensing uses some salient parameters: sensitivity, selectivity, response time, recovery time, stability, detection limit, and optimal working temperature [147]. As shown in Fig. 2.1, the curve between response and recovery time represents considerable information. The constant part depicts the approximate steady-state values. Since the inception of MOX gas sensors, only steady-state responses have been used to classify and quantify the gases for decades. Also, the steady-state responses are popularly knowns as the gold standard in gas sensing [86, 148]. However, the approximated steady-state is achieved after a long

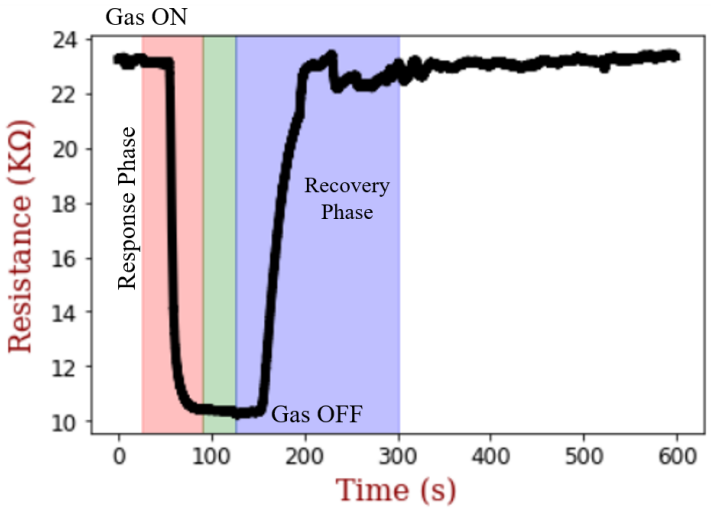

FIGURE 2.1: Response of a MOX (TGS2610) Gas Sensor for Ethylene at 125 ppm

time, which constrains the use of steady-state responses to classify the gases in the real-time scenario. This constraint opens the door to using dynamic responses. After decades, the dynamic responses had come into the picture. Since the dynamic responses are achieved by sampling the curve from response to recovery time, it results in high-dimensional data. When the steady-state responses are used, processing the data with less computational power is possible. But the sampled dynamic responses contain not only transient responses but also steady-state responses. Such high dimensional dataset requires additional pre-processing steps to prepare the data for classifying gases/odors.

2.1.4 Steady-State and Sampled Dynamic/Transient Responses

Steady-State Response: As indicated by the steady-state response, a static value is obtained when the gas sensor reaches a steady state. Hence, each gas exposure at a particular concentration results in a single-valued data point. The gas concentration is varied to obtain a dataset consisting of more steady-state responses for gas analysis. When instead of a single gas sensor, a gas sensor array is used for gas

sensing, it provides a steady-state response consisting of data points equal to the number of gas sensors used in the array. Thus, obtained steady-state responses are information-rich and better characterize the sensed gas or gases. As shown in Fig. 2.2, for a single gas sensor single steady-state region provides one static value, while using an array with four gas sensors, we can achieve four steady-state values making the resulting data vector more comprehensive.

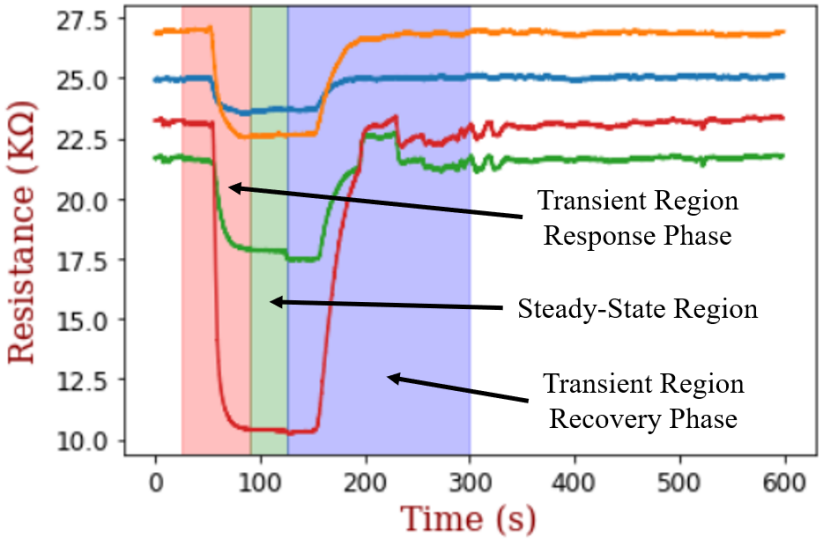

FIGURE 2.2: Steady-State and Transient Regions of Sensor Array Responses Captured for Carbon Monoxide and Ethylene

Sampled Dynamic/Transient Response: It has also been shown in Fig. 2.2 that there are two transient regions, one in the response phase and the second in the recovery phase. Between these transient regions, the steady-state is found. While using the gas sensor responses for real-time gas classification, there are three possibilities to collect the datasets:

1. When the transient found in the response phase is sampled.

2. When the responses are sampled, which include transient in the response phase and static responses.

3. The complete responses are sampled from transient in the responses phase to transient in the recovery phase.

Thus obtained datasets are high-dimensional and complex. Also, the considerable duration and high sampling rate can further enlarge the data making it vast and complicated. While using the traditional way to classify the gases using such dynamic responses, they can not be applied without proper data pre-processing that requires additional computational power. Hence, the transient or mixed responses can be used to classify the gases in a real-time scenario but at the expense of computational power. Also, the dynamic responses have extensive drift compared to the steady-state responses.

We have used four different datasets obtained from four different array of MOX gas sensors to demonstrate the presented works in this thesis. Three of them are publicly available to academic purposes while the fourth was recorded in the laboratory of our own institutes's department. Further, these datasets have been discussed in detail to understand the modality of each dataset.

2.1.5 Dataset-1

This extensive dataset was recorded to study the most challenging drift issue in gas sensing. It was collected over the span of 36 months for six gases/odors, viz., ethanol, ethylene, ammonia, acetone, acetaldehyde, and toluene. A gas sensor array consisting of sixteen gas sensor elements was used for the purpose mentioned above under a well-tuned working ambiance. The authors have used an ensemble learning-based machine learning approach to demonstrate the classification of considered gases/odors. In this experiment, the authors use four categories of commercially available Taguchi Gas Sensors, viz., TGS 2600, TGS 2602, TGS 2610, and TGS

2620. There have been used four gas sensors for each category. The corresponding gas sensor array was placed in a chamber of 60ml volumetric capacity where the desired gas/odor was exposed using a gas delivery system to record the responses. With this system, each response is obtained with sixteen time-series data vectors. The complete dataset consists of 13910 samples captured over the span of three years (see Table 2.1). The data acquisition takes 300s to complete measurement considering the gas injection phase for 100s and the recovery phase for 200s [86].

2.1.6 Dataset-2

This dataset was recorded to study the calibration transfer between gas sensor arrays. Four gases/odors, viz., ethanol, ethylene, carbon monoxide, and methane, were exposed to five replicas of a gas sensor array consisting of eight gas sensor elements for data collection. They have applied the master-slave technique to study the transferability of the calibration models. This dataset was made publicly available for academic purposes along with the publication of the above hypothesis. This dataset consists of 640 samples collected over the span of 22 days, exposing four gases mentioned above at ten different concentration levels to the five replicas of the considered gas sensor array (see Table 2.2 & Table 2.3). In this experiment, the authors use four categories of commercially available Taguchi Gas Sensors, viz., TGS 2602, TGS 2610, TGS 2611, and TGS 2612. There have been used two gas sensors for each category at the two different levels of the heater voltage [149].

2.1.7 Dataset-3

This dataset was captured while designing a gas sensing system for bioinspired early detection. The corresponding system comprises sixteen MOX gas sensor elements

Table 2.1: Related Content of Dataset-1.

Gas/Odor or Analyte	Discrete Concentration Doses	Month-Number of Recorded Samples
Ethanol	10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 100, 110, 120, 125, 130, 140, 150, 160, 170, 175, 180, 190, 200, 210, 220, 225, 230, 240, 250, 275, 500, 600.	1-84, 2-6, 3-10, 4-82, 9-11, 10-1, 11-360, 13-5, 14-52, 15-12, 16-28, 19-264, 20-250, 21-649, 23-30, 30-61, 36-600.
Ethylene	10, 20, 25, 30, 35, 40, 50, 60, 70, 75, 90, 100, 110, 120, 125, 130, 140, 150, 160, 170, 175, 180, 190, 200, 210, 220, 225, 230, 240, 250, 275, 300.	1-88, 2-10, 3-140, 4-170, 8-20, 9-4, 11-146, 12-334, 13-10, 14-43, 16-40, 17-20, 18-3, 19-100, 20-451, 21-662, 23-30, 30-55, 36-600.
Ammonia	50, 60, 70, 75, 80, 90, 100, 110, 120, 125, 130, 140, 150, 160, 170, 175, 180, 190, 200, 210, 220, 225, 230, 240, 250, 260, 270, 275, 280, 290, 300, 350, 400, 450, 500, 600, 700, 750, 800, 900, 950, 1000.	1-76, 2-7, 10-100, 13-216, 15-12, 16-20, 19-110, 21-360, 22-25, 23-15, 30-100, 36-600.
Acetone	12, 25, 38, 50, 60, 62, 70, 75, 80, 88, 90, 100, 110, 120, 125, 130, 140, 150, 170, 175, 180, 190, 200, 210, 220, 225, 230, 240, 250, 260, 270, 275, 280, 290, 300, 350, 400, 450, 500, 1000.	2-70, 3-7, 10-525, 13-275, 15-12, 16-63, 19-140, 20-466, 21-630, 22-123, 23-20, 24-28, 30-50, 36-600.
Acet-aldehyde	5, 10, 13, 20, 25, 30, 35, 40, 45, 50, 60, 70, 75, 80, 90, 100, 120, 125, 130, 140, 150, 160, 170, 175, 180, 190, 200, 210, 220, 225, 230, 240, 250, 275, 300, 500.	2-30, 4-4, 10-105, 12-192, 13-48, 14-18, 15-12, 16-46, 19-29, 21-744, 22-15, 23-18, 24-25, 30-50, 36-600.
Toluene	10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100.	2-74, 4-5, 19-9, 20-458, 21-568, 23-18, 24-1, 30-100, 36-600.

and had been tested using two gases/odors, viz., ethanol, acetone, and their binary mixture for the proposed hypothesis. This dataset was made publicly available for academic purposes along with the publication of the above experimentation. This dataset consists of 58 samples captured while exposing the gases/odors mentioned

Table 2.2: Related Content of Dataset-2 (1/2).

Gas/Odor or Analyte	Discrete Concentration Doses	Number of Recorded Samples
Ethanol	12.5, 25.0, 37.5, 50.0, 62.5, 75.0, 87.5, 100.0, 112.5, 125.0	160
Ethylene	12.5, 25.0, 37.5, 50.0, 62.5, 75.0, 87.5, 100.0, 112.5, 125.0	160
Carbon Monoxide	25.0, 50.0, 75.0, 100.0, 125.0, 150.0, 175.0, 200.0, 225.0, 250.0	160
Methane	25.0, 50.0, 75.0, 100.0, 125.0, 150.0, 175.0, 200.0, 225.0, 250.0	160

Table 2.3: Related Content of Dataset-2 (2/2).

Days for Testing the Gas Sensor Array Replica (Within Duration of 22 Days)	Replica Number
$4^{th}, 10^{th}, 15^{th}, 21^{st}$	1
$1^{st}, 7^{th}, 11^{th}, 16^{st}$	2
$2^{nd}, 8^{th}, 14^{th}, 17^{st}$	3
$3^{rd}, 9^{th}$	4
$18^{th}, 22^{nd}$	5

above and their binary mixture at different concentration levels. In this experiment, the authors use five categories of commercially available Taguchi Gas Sensors, viz., TGS 2600, TGS 2602, TGS 2610, TGS 2611, and TGS 2620. There have been used 2 (R4, R16), 2 (R3, R11), 8 (R1, R2, R5, R7, R9, R13, R14, R15), 2 (R6, R12), and 2 (R8, R10) gas sensors for the respective categories at the two different voltage levels and load resistors [150].

2.1.8 Dataset-4

This dataset was captured using an integrated gas sensor array of four gas sensing elements. It was fabricated using thick-film technology and aimed to study the sensitivity and response time. The sensing elements mentioned above were fabricated on a base material of tin oxide using the doping of cadmium sulfide, molybdenum

oxide, and zinc oxide. The fourth one was used as undoped tin oxide. The sensor characteristics were acquired using this integrated gas sensor array which was exposed to four gases/odors, acetone, carbon tetrachloride, ethyl methyl ketone, and xylene [151]. This dataset consists of two sets having 42 and 16 samples, respectively. The extraction of these samples is elaborated by Rajput et al. (2010) lucidly [152].

2.2 Methods

The data analytics for classifying and quantifying the gases/odors employ various pattern recognition techniques. However, several traditional pattern recognition techniques are promising in specific applications dealing with discriminating and estimating gases/odors. But the dominance of neural networks is still sustained for each task at hand in the context of gas sensor response analysis due to their diverse suitability. Moreover, neural networks have few limitations compared to their peers for the same purposes.

2.2.1 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are the traditional ones that use only fully connected layers or dense layers. These layers require input in the form of one-dimensional vectors. Due to full connection, these layers generate huge trainable parameters, which make the computational process sluggish. Moreover, such networks learn only linear variance-based salient discriminable features inherently confined in the 1D vectors. A typical ANN for classifying the gases/odors has been shown in Fig. 2.3. Each node in the input layer represents the response vector

obtained through the gas sensor array. At the same time, each node in the output layer results in a probability estimate corresponding to the target. For the final classification of the input vector, the target with the highest probability estimate is considered the desired target. As shown in Fig. 2.3, the layer between input and output is popularly known as the hidden layer, which may be used in numbers as desired but results in the complex architecture. Each hidden layer node represents a mathematical function called "perceptron." Hence, each fully connected layer (hidden layer) is a parallel arrangement of perceptrons. Thereby, ANNs are also called multilayer perceptron (MLP) networks. A schematic diagram of a perceptron is shown in Fig. 2.4.

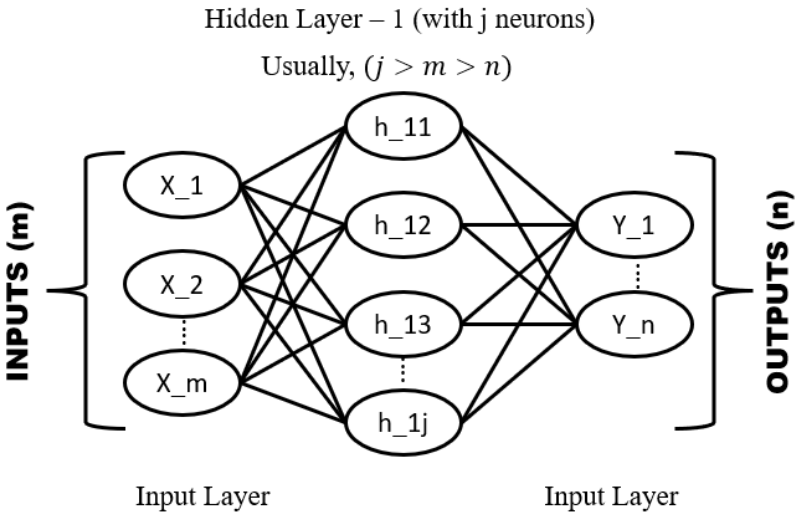

FIGURE 2.3: A General Architecture of Artificial Neural Network

Except for input and output value, a perceptron or neuron or the unit cell of neural networks has three parameters: Weight, Bias, and Activation Function. **Weight** is the parameter of amplification or attenuation to the input value. Each weight corresponding to the inputs is generated randomly following some stochastic process. With these weights, the weighted sum of the inputs is used as the input to a bounded input bounded output (BIBO) **Activation Function** that provides the

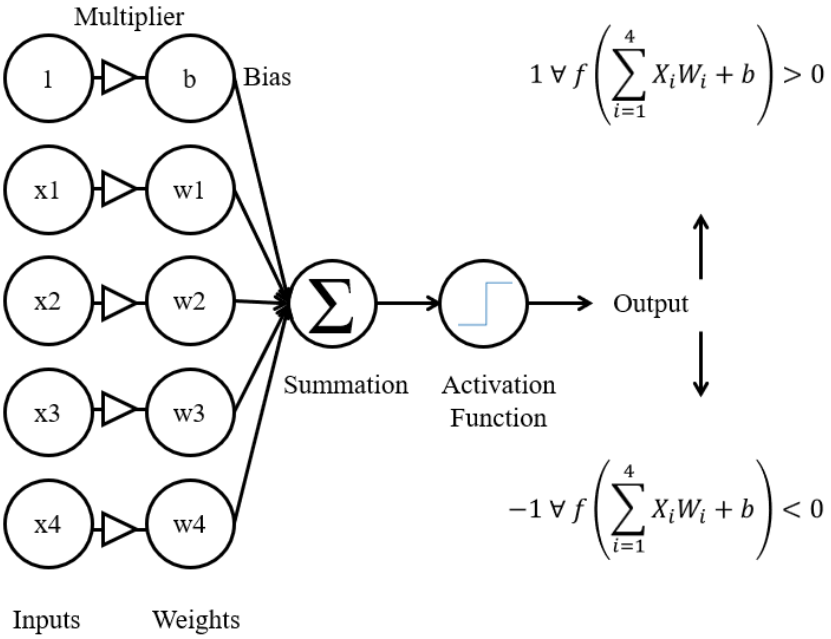

FIGURE 2.4: Perceptron Architecture

final output of the perceptron. Activation functions are used to learn the nonlinear relationship between the input and output. However, no activation function is ideal to use for every problem. It must have some properties according to the task. Consider the case when all inputs become zero due to some noise effect or other inconvenience; consequently, the weighted sum also becomes zero. Several activation functions cannot be defined explicitly for zero input value in such cases. An offset value (constant) is added to the weighted sum to avoid the situation when an activation function gets zero input value. This value is popularly known as the **Bias** and is typically taken 1 (by default). Several activation functions are frequently used by researchers, which are as follows:

One of the primitive activation functions is a linear function. Because of its constant gradient, it is independent of inputs and cannot utilize the updated weights and biases to improve the performance. Thus, it is not suitable to use in hidden layers. However, many single-layered ANNs perform satisfactorily by using it as an

activation function. Usually, the linear activation functions are used in output layers where the continuous outputs are estimated. The most widely used activation function is sigmoid. It has several advantages due to being continuously differentiable and one of the non-linear functions. But it has the problem of vanishing gradient. Also, the range of sigmoid functions is [0,1], which is not centered on the origin. Due to its asymmetric nature, it can only provide positive values. In contrast to these issues, it is frequently used in output layers for binary classification problems. Moreover, the hyperbolic tangent (*tanh*) is also a significant activation function that resolves the asymmetricity issue in the sigmoid activation function. It ranges from -1 to +1 and is centered on the origin. This property makes it suitable for optimization and is frequently used in hidden layers. Although, it also faces vanishing gradient issues. The most widely used activation function for hidden layers is rectified linear units (ReLU). It is computationally efficient compared to the discussed activation functions. As mentioned earlier, the sigmoid function is used as the output activation function for binary classification; a quite similar activation function is the softmax activation function, which is essentially used as the output activation function for multiclass classification problems. It generates the probability estimates for each input corresponding to each target. Furthermore, various modifications of some of these activation functions also exist in the literature to overcome the limitations of discussed activation functions.

2.2.2 ANN Compatibility with Array Response Modalities

As discussed earlier, if the input data vectors are obtained using the steady-state responses, each vector consists of data points equal to the number of gas sensors used in the array. Typically, a gas sensor array may have 4 to 16 gas sensors, as observed by reviewing the literature. However, no thumb rule is existed in the literature to

use an optimal number of gas sensors in the array. With this fact, if a gas sensor array has four gas sensors results in input data vectors having only four data points. These small data vectors require a simple ANN to classify the target gases/odors.

On the other hand, while using the transient responses, the resulted input data vectors have huge sizes depending on the sampling rate and duration. Such data vectors are significantly correlated and cannot be recommended to use directly to save computational power and time. Because processing the raw input data vectors as it is, requires a highly complex ANN architecture. Also, raw input data vectors may overfit the training model. Several pre-processing methods are applied to raw data vectors in such cases for various purposes: dimensionality reduction, feature selection, feature extraction, etc.

With the evolution of convolution and pooling layers, ANNs transformed into the new paradigm called deep neural networks (DNNs). In neural networks, the word "deep" used for two different inferences: firstly, for indicating more than two hidden layers in ANNs; and secondly, for demonstrating the use of convolutional and pooling layers. The ANNs involving convolution and pooling layers are popularly known as convolutional neural networks (CNNs). The CNNs have set a benchmark in the field of neural networks. They are worldwide famous for image classification problems. The forthcoming section discusses how the researchers have started to use CNNs in the field of gas sensing. Also, we have discussed our motivation to use CNNs for gas sensing in a novel way that has never been used before.

2.3 Technology Background

Except for CNNs, traditional pattern recognition techniques, including ANNs, require 1D data vectors for classifying the targets. Moreover, for providing better

performance, these are highly dependent on the handcrafted (manually extracted) features. These features subsequently face the problem of feature selection or feature ranking. Determining relevant features also adds a computational burden to accomplishing the task at hand. A CNN is a panacea for all such problems where the extraction of features is handled manually. CNN's capability of automatically extracting high-level features makes it a superperformer. As quoted earlier, it is a well-known worldwide technique that got attention from the image classification tasks. An image is a 2D structure (binary and grayscale images) or a 3D structure (color, multispectral, and hyperspectral images) of the data. For example, a pioneer CNN has been used to classify the MNIST data [153] which consists of grayscale images of handwritten digits. On the other hand, CNN has been utilized to classify Hyperspectral Images [154].

The pioneered CNN is also called a 2D-CNN due to its two-dimensional operationality for convolution and pooling operation. However, using the contextual outlines, its one-dimensional and three-dimensional forms developed later. Furthermore, every version of CNN has significance depending on the application, but 2D-CNNs have been more frequently customized to utilize their importance in diverse areas. For illustration, the first use of a 2D-CNN in the gas sensing area was published in 2017, and the related works in the gas sensing area witness that 2D-CNNs have been efficiently used for various purposes. Before discussing the use of a 2D-CNN for gas sensing, it would be worth detailing a general architecture of a 2D-CNN first.

2.3.1 A General Architecture of 2D-CNN

Originally, CNNs were designed to extract the salient features from the 2D input vectors while dealing with classification problems. They are specialized architectures of neural networks. Compared to traditional neural networks CNNs are computationally efficient. The general and most straightforward architecture of 2D-CNN has been shown in Fig 2.5. It has two specialized components in the form of convolution and pooling layers, making it different from traditional architectures.

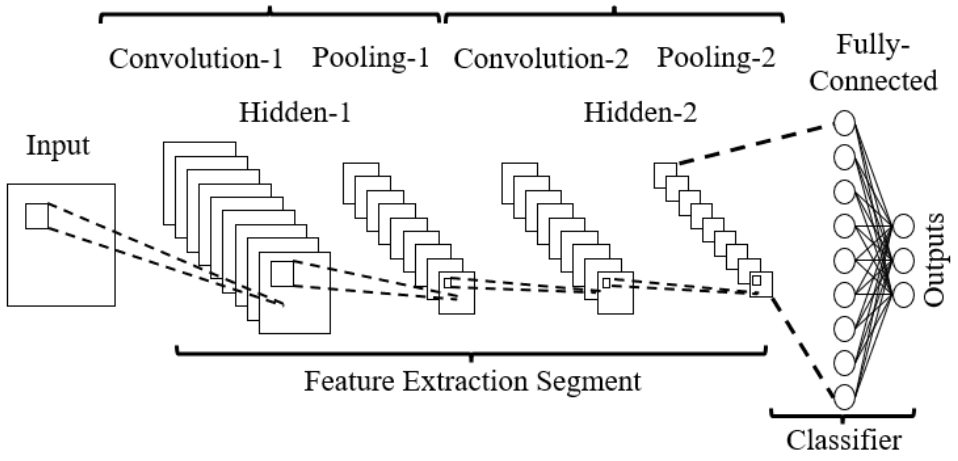

FIGURE 2.5: A Simplest Architecture of 2D-CNN

The convolution layer plays a role in implementing certain operations to extract the features similar to perceptron in ANNs. In convolution operation, a filter or kernel (a set of weights) operates on the input data covering all corresponding receptive fields to generate the respective feature maps. A schematic diagram performing a 2D convolution operation is shown in Fig. 2.6.

While working on images, the receptive fields are highly correlated. Therefore pooling layers are used to down-sample the resulting feature maps, which reduces the dimensionality and thereby the trainable parameter. In this way, pooling layers make the architecture computationally efficient. Pooling is performed to keep

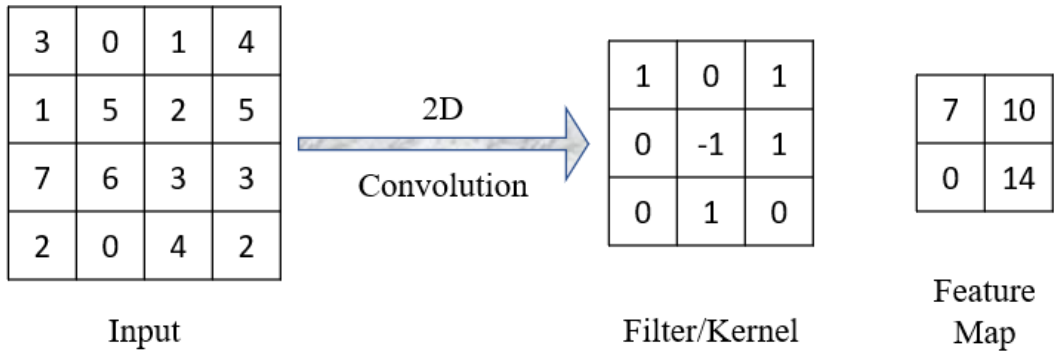

FIGURE 2.6: 2D Convolution Operation

the maximum, average, and minimum values of the pooled window. A schematic diagram performing a 2D max-pooling operation is shown in Fig. 2.7.

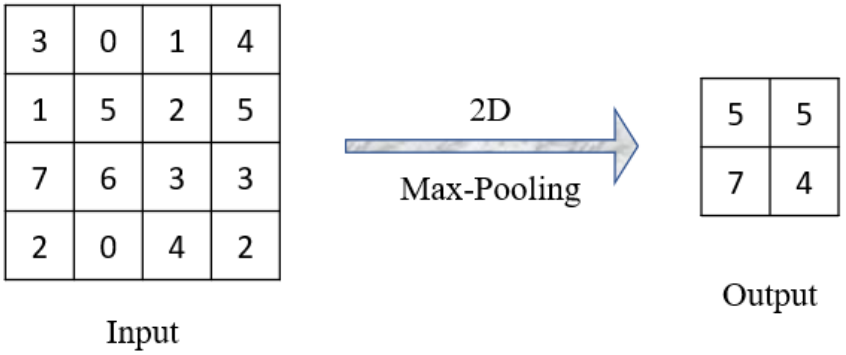

FIGURE 2.7: 2D Max-Pooling Operation

2.3.2 Motivation for Using 2D-CNN for Gas Classification

As evident from the discussion done till now, 2D-CNN was developed to be operated on 2D data structures (especially images). With the evolution to date, 2D-CNN is applied to images in two styles. In the first case, a whole image is used as the input; in the second case, a group of pixels (patch) is used as the input. For illustration, see Fig. 2.8.

Using the second case, we have used 2D-CNN in various works, which have been published outside the scope of the thesis presented [155–158]. In these works,

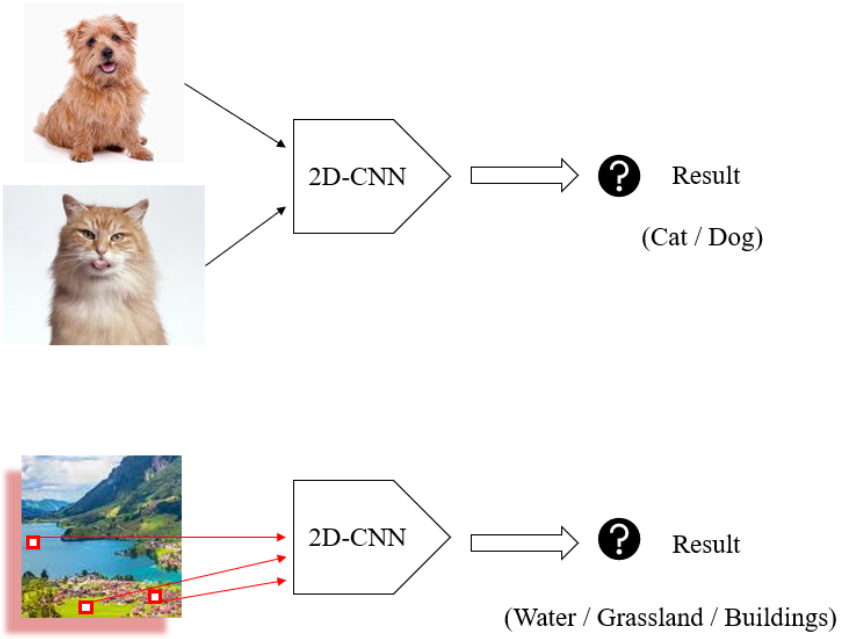

FIGURE 2.8: 2D-CNN for Image Classification

a group of pixels (patch) is used as the input to classify the targets. These small patches are squared 2D data structures and have shapes such as (5×5), (7×7), (9×9), and so on [154]. We have inferred from here that a minor patch having size (5×5) can be used as the input to the 2D-CNN. As quoted earlier, a gas sensor array typically has 4 to 16 gas sensor elements, resulting in 4 to 16 data points for each sample recorded for each observation. Thus, the obtained response vectors can be represented in 2D-squared structures having shapes ranging from (2×2) to (4×4). With this size, our response vectors are still incompatible with the 2D-CNN. To overcome this constraint, we have successfully proposed specialized data transformation techniques. However, 2D-CNNs have already been applied to gas sensing since 2017. But authors/researchers have used highly complex architectures for gas sensing. Such highly complex architectures are worth using while working with actual image datasets. In contrast, in gas sensing, the data is transformed to represent in a 2D format to mimic the actual image synthetically. Also, the found 2D-CNNs for gas classification to date can be applied to only dynamic responses; they

fail when only the steady-state responses are available to classify the gases/odors. We have also overcome this constraint.

2.4 Conclusion

In the gas sensing area, a variety of technology is available to design the gas sensor depending on several gas detection principles. Broadly, the researchers have categorized the gas sensors into two types. Out of those, one kind of gas sensor detects the gases based on the variation in electrical properties of the sensing materials, and another kind of gas sensor utilizes non-electrical properties for gas sensing. However, MOX gas sensors that fall under the first category are much more prevalent in the gas sensing research community. It doesn't mean that the MOX gas sensors serve accurately everywhere. The MOX gas sensors have several merits: low cost, fast response, long lifetime, suitable for miniaturization, etc. Also, it has several demerits: poor selectivity, drifted responses, highly dependent on temperature and humidity, etc. But the synergy of MOX gas sensor-based array and advanced pattern recognition techniques are setting benchmarks in the areas of gas sensing. Almost all pattern recognition techniques have been used to classify the gases/odors using responses obtained from MOX gas sensors. However, neural networks have played a dominant role for decades. Traditionally, ANNs are used to classify the gases/odors, but such networks have only fully-connected layers, which provide considerable parameters to train the network. The evolution of advanced layers for convolution and pooling made the neural networks less complex. These layers offer parameter sharing and sparse connections to the traditional architectures. Moreover, ANNs are applied to manually processed data for feature enhancement. But the convolutional

neural networks that use the convolution and pooling layers can extract high-level features automatically to provide outperforming results.

With this inference, we have customized the convolutional neural networks well-known for image classification for gas sensing. Since gas sensor responses are inherently available in 1D vectors, we have proposed specialized data transformation techniques that make the gas sensor responses compatible with 2D-CNN. Also, we have developed a simpler 2D-CNN capable of discriminating gases/odors using the gas sensor array's dynamic and static responses. In the customized CNN, we do not use pooling layers since they necessarily serve the purpose only while using the CNNs for image classification. However, pooling layers may utilize their significance for highly complex and highly correlated non-imaging datasets. With our customized CNN, we have novely contribute for application of CNN for classification of gases/odors independent of gas sensor array responses modalities. Also, we have propose the optimization approach for gas sensor nodes suitable for resource-constraint environment. Moreover, we have devised a hybrid CNN architecture that can compensate the drift effects without using additional statistical algorithm for drift correction.