E-ISSN No : 2581-6195 | Volume : 9 | Special Issue | Aug 2026

MACHINE LEARNING ASSISTED OPTIMIZATION OF OXIDE NANOSTRUCTURES FOR RAPID AND SELECTIVE MULTI-GAS SENSING LIKE SnO_2, Nio, Nb_2O_3 or Nb_2O_5,ETC.

Access Online

Article DOI:
10.5281/zenodo.22765774

Sunny Sharma

Research Scholar,
Department of Physics,
Phonics University,
Roorkee

How to Cite:
Sunny Sharma (2026), Machine Learning Assisted Optimization of Oxide Nanostructures for Rapid and Selective Multi-Gas Sensing Like SnO_2, Nio, Nb_2O_3 or Nb_2O_5,etc., International Educational Journal of Science & Engineering (IEJSE),Vol: 9, Special Issue, Aug 2026 79-90

ABSTRACT

Tin (IV) oxide (SnO_2), Nickel (II) oxide (NiO), Niobium(V) oxide (Nb_2O_5) and Niobium (III) oxide (Nb_2O_3) are metal oxide semiconductor (MOS) nanostructures, which have thus far been at the heart of chemiresistive gas sensing due to their adjustable surface chemistry, high sensitivity, and ability to be microfabricated. Although heavily material-engineered, conventional MOS sensors typically lack cross-sensitivity, have low levels of inherent selectivity, have sluggish response-recovery kinetics, and drift in multi-gas conditions. Latest progress has demonstrated that the combination of machine learning (ML) and oxide nanostructured sensor arrays can greatly improve the performance of gas discrimination on complex atmospheres. As an example, more than 99 percent accuracy in classification of volatile organic compounds based on MOS sensor arrays when analyzed with a Random Forest and k-Nearest Neighbor algorithm, and deep learning methods have allowed detection of dangerous gases in real-time with very high accuracy.

This is a critical review of the role of ML-assisted strategies in optimization in the form of supervised learning, feature engineering, dimensionality reduction and neural network structures in enhancing selectivity, drift compensation, and morphology-performance relationships in MOS sensors. This discussion points at the two issues, which include the lack of data and the interpretation of models, and research and development perspectives of self-calibration edge-AI-powered sensor arrays and physics-informed ML models transforming future gas sensing technologies.

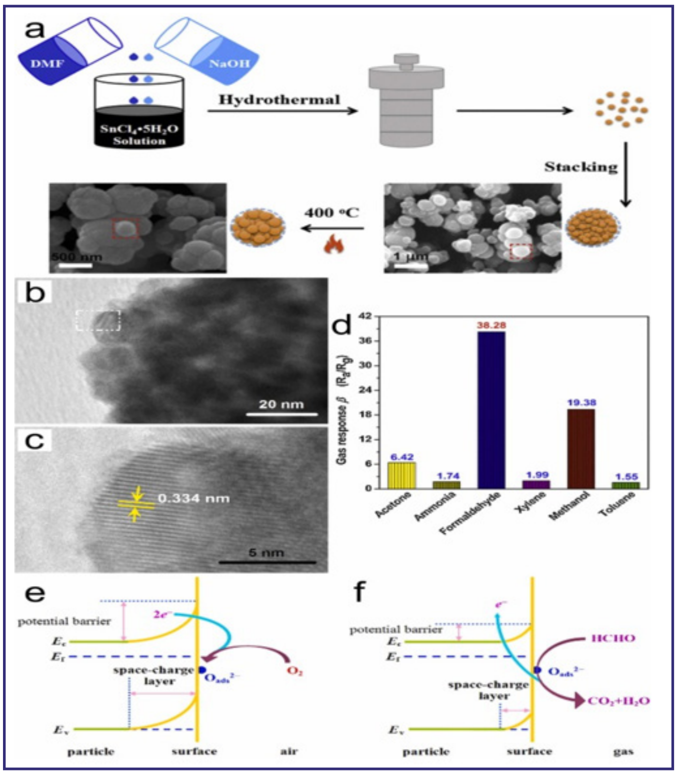
KEYWORDS: Metal Oxide Semiconductors (MOS), Chemiresistive Gas Sensing, Machine Learning-assisted Gas Sensors, Sensor Drift Compensation, Gas Selectivity Enhancement, Edge AI Sensor Arrays

1. INTRODUCTION

The rapid and selective detection of hazardous gases is essential for environmental monitoring, industrial safety, smart infrastructure, and medical diagnostics. Toxic gases such as carbon monoxide (CO), nitrogen dioxide (NO_2), ammonia (NH_3), and volatile organic compounds (VOCs) present severe health and environmental risks even at trace concentrations, requiring sensors capable of low detection limits and high discrimination capabilities. Metal oxide semiconductor (MOS) sensors have remained a dominant technology due to their cost-effectiveness, chemical stability, and ease of integration (Han et al., 2025). The fundamental gas sensing mechanism in MOS devices is based on adsorption-induced modulation of electrical resistance: gas molecules interact with adsorbed oxygen species on the oxide surface, leading to charge transfer that alters the depletion layer and measurable conductivity changes (Han et al., 2025). Nano structuring MOS materials — such as forming nanowires, nanosheets, or hierarchical porous morphologies —

Research Paper

increases active surface area and enhances sensitivity by facilitating surface reactions and gas diffusion.



Source:* *Masuda, Y. (2022). *Recent advances in SnO_2 nanostructure based gas sensors.* ***Sensors and Actuators B: Chemical,*** *364, 131876.*

However, even with careful tuning of morphology and surface chemistry, conventional MOS sensors face limitations that become particularly problematic in multi-gas environments. Cross-sensitivity arises when different analytes produce overlapping response signatures, while temperature variations and humidity can induce drift and baseline instability. These issues make reliable multi-gas discrimination difficult with single sensors or simple arrays of differently doped materials. In addition, as reported by Han et al. (2025), high sensitivity often comes at the cost of increased false positive rates under mixed gas conditions, highlighting intrinsic selectivity challenges.

To address these limitations, machine learning has emerged as a powerful data-driven paradigm capable of analyzing complex sensor response patterns across time, frequency, and multi-sensor datasets. Supervised learning techniques such as random forests and k-nearest neighbors have been successfully applied to classify gases based on dynamic response fingerprints, achieving high accuracy even in mixtures with significant cross-interference (Singh et al., 2024). Deep learning models trained on temperature-modulated or pulsed response features have demonstrated rapid and precise identification of multiple target gases using either sensor arrays or single MEMS-based devices (Lee et al., 2025; Kim et al., 2024). These methods leverage high-dimensional feature spaces to distinguish subtle nonlinear response characteristics that traditional linear calibration methods cannot resolve. Moreover, recent work has begun to explore ML-based descriptor discovery for doped oxide gas sensing, identifying key physicochemical features that govern sensor behavior (Su et al., 2025).

2. OXIDE NANOSTRUCTURES FOR ADVANCED GAS SENSING

2.1 SnO_2 Nanostructures: Surface Engineering and Performance Evolution

Tin (IV) oxide is the most widely studied n-type Nano semiconductor in chemiresistive gas sensing as it exhibits high hereditary mobility, thermal stability and programmable oxygen vacancy chemistry. More recent studies have been done between 2020 and 2025 on defect engineering, heterostructure formation, and machine integration of microelectromechanical system (MEMS) to improve selectivity and low operating temperatures. Vasantharao and Dhar (2025) have logically shown that SnO2 nanowires and hierarchical porous nanoflowers are more sensitive than compact thin films since the crystal size is close to Debye length, resulting in complete depletion in ambient conditions. This complete depletion regime optimizes modulation of resistance when the gas adsorbs and this increases the response magnitude.

Electrospinning, hydrothermal growth, and atomic layer deposition were some of the advanced synthesis strategies which allowed the control over grain boundary density and surface-active sites with high precision. Functionalization of noble metals, especially Pt and Pd nanoparticles, will increase catalytic dissociation of target gases, which makes surface reactions faster and reduces the activation energy barrier (Han et al., 2025). Moreover, it was demonstrated that temperature-controlled operation and ML-based feature extraction would achieve a big increase in the capability of gas discrimination in SnO2 arrays (Sensors & Actuators B, 2025). The effects of the interacial band bending associated with the introduction of heterojunction systems like SnO2-TiO2 composites alter the electron transporting pathways to enhance both the sensitivity and speed of response. Non-linear signal behavior that cannot be completely explained by classical analytical calibration is however introduced by increased complexity of these nanostructures hence the need to incorporate machine learning (Han et al. 2025).

Table 1: ML-Assisted Oxide Nanostructure Gas Sensors

Study / Material System	Oxide Nano-structure	Target Gases	ML Model / Algorithm	Key Outcome	Source
Mesoporous SnO_2 QDs	SnO_2 quantum dots (mesoporous)	H_2S	XGBoost feature importance	Identified optimal microstructure parameters; enhanced response ~30× vs non-porous	Liu et al. – *Chemosensors*
Sb-doped SnO_2 array	Sb-SnO_2 films	CO, NH_3, H_2, NO_2	ML classification (fingerprints)	Near-perfect multi-gas classification from single material	*Journal of Alloys and Compounds*
SnO_2–TiO_2 hetero-junction	SnO_2/TiO_2 on MXene	H_2 & NH_3	ML enhanced selectivity	High selectivity and sensitivity in mixed gases	Sensors & Actuators B
NiO-doped SnO_2 + PCA	NiO-SnO_2	H_2 & NH_3	PCA + ML classification	100 % identification accuracy	*J. Semiconductors*

2.2 NiO Nanostructures: p-Type Sensing and Heterojunction Synergy

Nickel (II) oxide has seen a resurgence in the 2020-2025, especially due to the complementary p-type conduction mechanism and high sensitivity to the reducing gases, e.g. hydrogen and ammonia. Kaur (2025) has indicated that NiO nanoflakes and mesoporous thin films have a high surface reactivity and hole transport due to a large number of Ni3+ defects and oxygen vacancies. The p-type response mechanism entails recombination of holes with electrons produced in the course of gas-surface reactions leading to the resistance increment on exposure to reducing gases.

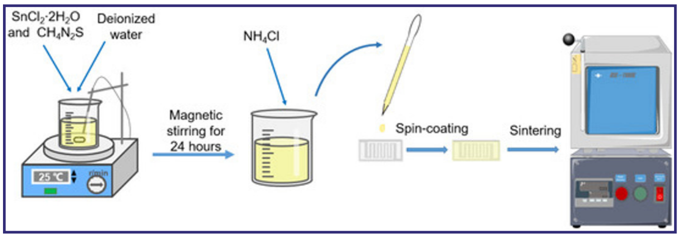
In recent research, p-n heterojunction structures based on NiO in the form of NiOxSnO2 or ZnO, in which interfacial depletion layers are used to enhance resistance modulation. The synergistic properties in these heterostructures because of the inherent electric fields at the interface enhance the efficiency of charge carrier separation and amplitude of response (Han et al., 2025). However, the dynamics of the reaction of NiO-based systems is very sensitive to humidity and ambient oxygen concentration. It has thus been observed that data-driven modeling methods have been used more to balance environmental variability and enhance the accuracy of gas classification (Lee et al., 2025). NiO nanostructures combined with ML algorithms have shown significant enhancement of multi-gas selectivity, especially with transient dynamic properties being taken into consideration in data training.

2.3 Nb_2O_5 and Nb_2O_3: Emerging Niobium-Based Oxide Sensors

Recently, Niobium(V) oxide is being proposed as a potential candidate of VOC and humidity sensor as it is highly chemically stable and reacts well with polar molecules. The Nb2O5 mesoporous films between 2020 and 2024 reported that the films have an increased surface hydroxyl density, which allows the conductivity to be controlled by adsorption. It has several crystalline phases that enable customizable electronic structures and defect states to be produced that can be optimally designed to enhance sensing selectivity. Nonetheless, Nb2O5 frequently has lower response kinetics than SnO2 meaning that it requires algorithmic compensation approaches to real-time identification.

The lower oxidation state Niobium (III) oxide has a better intrinsic electronic conductivity owing to a higher oxygen vacancy concentration. Oxygen-deficient Nb2O3 thin films produced by reactive sputtering have been investigated recently, with better response to the reduction of gases at moderate temperatures. Nevertheless, the change in stoichiometry has a big influence on the baseline resistance and reproducibility. Such variability poses a problem to the calibration process, which can be addressed using ML-based regression models that can learn nonlinear structure response correlations (Han et al. 2025).

Fig. 1: The preparation process of MPTD gas sensors.



2.4 Structure–Property Relationships and Nonlinear Behavior

The complex dependence between the crystallite size, surface defect density, grain boundary potential barriers, and the diffusion pathways of the gas phase are the controlling factors of the sensing performance of oxide nanostructures. Larger particles are more sensitive to smaller particle sizes, but a further reduction in particle size will result in lack of stability and high noise. The influence of grain boundary barriers is that they change the potential height with exposure in the presence of gas. Porosity increases the rate of diffusion of gases but can bring about mechanical weakness. The combination of these interdependent factors produces a high dimensional parameter space which is hard to optimize solely based on empirical experimentation.

Recent works (Han et al., 2025; Sensors and Actuators B, 2025) also make it clear that multidimensional response signatures are generated in the nonlinear interactions between morphology, defect chemistry, operating temperature, and environmental conditions. The dependence is not reflected by conventional linear calibration. As an alternative, machine learning structures offer a powerful approach to nonlinear correlations of input variables (structural parameters, environmental conditions) and output performance measures (sensitivity, selectivity, response time). Therefore, the optimization of multi-gas sensors has become unthinkable without the help of ML in the translation of advanced nanostructures into dependable systems.

3. MACHINE LEARNING ALGORITHMS FOR MULTI-GAS CLASSIFICATION AND OPTIMIZATION

The development of machine learning architecture has largely contributed to the increase in the analytical performance of metal oxide semiconductor gas sensors since 2020 through rapid advancement of machine learning methodologies. With the nanostructured sensing materials like Tin (IV) oxide, Nickel (II) oxide, Niobium(V) oxide and Niobium (III) oxide giving rise to progressively complex and nonlinear response signatures classical calibration techniques, relying exclusively on steady-state resistance values, have become inadequate. Rather, greater use of multidimensional temporal, spectral, and cross-sensor data is harnessed by ML algorithms to allow proper classification of gases, estimate concentrations and correct drift under realistic operating conditions.

3.1 Data-Driven Feature Engineering for Metal Oxide Gas Sensor Arrays

Among the most significant applications of machine learning in the current gas sensing studies, the analysis and recognition of multidimensional response trends produced by metal oxide semiconductor (MOS) nanostructures can be listed. SnO2, NiO, Nb2O3 and Nb2O5 are some of these materials that show change in dynamic resistance with various gases. Such differences depend not only on the target gas but also the operating temperature, humidity, surface morphology, grain boundaries and catalytic additives. The raw electrical response of a sensor due to a combination of these various interacting parameters will often have hidden patterns that cannot be easily interpreted using the standard signal analysis tools.

Machine learning offers an effective platform which can be used to obtain meaningful information with these multifaceted datasets by feature engineering and pattern recognition. In MOS gas sensing, sensor response features may be based on a range of sensor response parameters such as transient response curves, response and recovery time constants, rate of resistance change, frequency-domain response and response patterns modulated by temperature. The combination of these features makes up a multi-dimensional dataset which is a characteristic of unique sensing behavior of each gas. With the help of such features, machine learning algorithms are able to recognize patterns which are related to certain chemical reactions that are present at the sensor surface.

As a case in point, nanostructured sensors of SnO2 usually exhibit different response kinetics when subjected to oxidizing versus reducing gases, e.g. NO 2 versus CO or H2. Likewise, p-type NiO sensors display the opposite trend in their resistance variation to n-type SnO2 and thus offer useful elements to heterogeneous arrays of sensors. When these materials are incorporated into an array of sensors, each sensing element has its response signature. These aggregate cues can be used to enhance gas discrimination with the help of machine learning models.

The techniques of extracting feature used in machine learning before the training of the models include principal component analysis (PCA), linear discriminant analysis (LDA), and time-series statistical descriptors. PCA will minimize the sensor dataset dimensionality and will retain the most relevant elements of variance in the sensor data. This allows the researchers to see the clusters that represent various gases and makes the classification easier. LDA also serves to optimize the separation of the different classes of gases and minimizes the separation in between the different classes themselves.

One more significant feature of feature engineering in gas sensing is the application of dynamic operation modes. Rather than using constant temperature sensors, temperature-modulated sensing strategies are now used to produce more rich datasets. In these systems, the sensor temperature of the MOS is periodically altered, and a series of patterns of response to a single exposure to the gas are obtained. With the help of this approach, much more information can be given to machine learning analysis. As an example, SnO2 nanostructures can exhibit a variety of adsorption kinetics at varying temperatures, which algorithms can be used to differentiate gases that would otherwise behave similarly at fixed conditions.

Other spectral analysis tools used to obtain frequency-domain features of sensor signals include fast Fourier transform (FFT) and wavelet transforms besides temperature modulation. The techniques transform the sensor functions over time into spectral components and identify patterns in those quantities, which are not observed in the time-varying sensor responses. These spectral features can then be used by machine learning algorithms in order to enhance the accuracy of classification.

The other design that seems encouraging is the combination of microelectromechanical systems (MEMS) technology with MOS nanostructures. MEMS-based gas sensors are able to provide high-resolution data of the temporal response because they are fast to react to thermal changes and consume low energy. This set of sensors together with machine learning can be used in quick and dependable sensing of numerous gases in miniature sensing frameworks.

The cost of using feature engineering methods is high because it greatly enhances the interpretability and predictive power of machine learning models in gas sensing. Using raw sensor responses to form organized datasets, researchers are able to obtain a clear correlation between the properties of nanostructures and the sensing performance. This is a method that eventually helps in the creation of smart sensor systems which can be used under the real-world conditions.

3.2 Machine Learning Guided Material Design and Sensor Performance Optimization

In addition to gas classification, the machine learning is also important in maximizing the structural and functional characteristics of oxide nanostructures that are utilized in gas sensing. SnO2, NiO, Nb2O3 and Nb2O5 are some materials, whose sensing performance is highly dependent on morphology, crystal structure, density of defects, and surface chemistry. Most of the classical experimental optimization techniques entail lengthy trial-and-error synthesis, which is time and resource consuming. Machine learning provides an alternative, which is data-driven, by being able to model structure-property relationships in a predictive manner.

Over the last few years, scholars have been using machine learning models more often to study experimental data on synthesis parameters, structural features, and sensing performance measures. The variables may be the concentration of the precursor, annealing temperature, doping elements, particle size, surface area, and the values of the gas response. The machine learning algorithms can be used to forecast the impact of any changes in the synthesis conditions on the sensing behavior of nanostructured materials by learning patterns using these datasets.

As an illustration, SnO_2 nanostructures may be prepared in different morphologies, which are nanoparticles, nanowires, nanosheets, and hollow nanospheres. The morphology provides varying surface area and electron transportation capacity, depending on each morphology with direct effect on gas sensing. Experimental data collected under many different synthesis methods can be studied by machine learning models and determine which structural parameters result in high sensitivity or short response times.

In a similar manner, NiO nanostructures are also being considered to detect gases like ethanol, acetone and hydrogen sulfide owing to their p-type semiconducting character. Doping with noble metals or heterojunctions with n-type oxides (e.g. SnO_2) can greatly improve the performance of NiO sensors. The efficacy of various doping methods may be assessed using machine learning to match the material composition with the indicators of sensing responses. This enables the researchers to find some of the best combinations between materials so as to maximize their selectivity to certain gases.

Nb_2O_5 and Nb_2O_3 are niobium oxide materials with promising gas sensing properties since these materials are highly chemically stable; in addition, they have tunable electronic properties. The sensing mechanisms of these materials are however, usually complicated in that they consist of multiple oxidation states and defect structures. Experimental data can be analyzed using machine learning models to identify the influences on gas adsorption behavior of changes in oxygen vacancies, crystal phases, and surface functional groups.

Optimization of sensors is another important benefit of machine learning as it can manage nonlinear dependencies among a number of variables. The performance of gas sensing is hardly defined by a single parameter but it is the product of interaction between structural properties, environmental conditions, and operating parameters. Random forest regression, gradient boosting, and neural networks are some of the algorithms that can capture all these nonlinear interactions and make correct predictions of sensor performance.

Besides materials optimization, machine learning may also be helpful in enhancing sensor stability and long-term reliability. Signal drift due to aging effects, contamination or environmental changes is one of the greatest issues in MOS gas sensors. Drift may cause substantial loss of precision of gas detecting systems over a period. The machine learning algorithms are able to correct the drift by frequently updating the calibration models using historical sensor measurements. This self-corrective learning skill allows the sensor system to be able to achieve reliable performance even when operating within different environments.

Moreover, machine learning applications are becoming common in order to optimize sensor operation parameters including heater temperature, sampling frequency and exposure time. Algorithms are able to give optimal operating conditions that give maximum classification accuracy and minimal detection time by processing high volumes of data gathered by sensor arrays. The method is especially useful when the portable, wearable gas sensing device is needed, and the energy consumption rate and responsiveness are the essential conditions.

Altogether, it can be stated that machine learning combined with nanostructured MOS gas sensors is a groundbreaking method of designing the next gas sensing technology. Through data-driven modeling along with materials engineering, scientists will be able to speed up the development of high-performance sensing materials and can create future intelligent sensor systems which will be able to detect multiple gases quickly, selectively, and reliably.

3.3 Feature Engineering and Dimensionality Reduction

In spite of deep learning, feature engineering is the key component of ML-assisted gas sensing (especially in systems with limited resources). The features that are commonly extracted are peak response amplitude, response and recovery time constants, area under the transient curve, derivative slopes, frequency spectrum components obtained by Fourier transforms and statistical characterizations of skewness and kurtosis. The temperature-modulated sensing adds more features based on frequencies, which is highly effective in improving discrimination capability.

Principal Component Analysis (PCA) is an example of dimensionality reduction techniques that remain an important tool of preprocessing. Even though PCA is an unsupervised algorithm, not explicitly involved in the classification process, it allows the visualization of the patterns of clustering and minimizes the computational load. A number of research works between 2020 and 2024 have shown that a combination of PCA and SVM or RF enhances the stability of classification by eliminating redundant or noisy features (Singh et al., 2024). LDA has also been used to maximize between-class variance in multi-gas data sets.

This further enhanced generalization of the models as the feature selection algorithms (recursive feature elimination and mutual information ranking) are integrated. Researchers have been able to minimize overfitting by removing irrelevant descriptors and the loss of accuracy is minimal especially in small datasets

Material & Morphology	Target Gas(es)	ML Algorithm / Data Analysis	Performance
SnO_2–TiO_2 heterojunction on MXene	H_2 & NH_3	SVM, ANN for mixed gas analysis	Enhanced selectivity and sensitivity; low H_2 detection (~200 ppb); significantly improved mixed-gas differentiation
Sb-doped SnO_2 sensor array	CO, NH_3, H_2, NO_2	ML fingerprint-based classification	Near-100% gas classification accuracy using doping-dependent response patterns
NiO-doped SnO_2 (transient response)	H_2 & NH_3	PCA + supervised ML classifier	Achieved 100% identification accuracy from single sensor transient features
SnO_2 nanosheet MEMS with pulsed heating	H_2, CO, NH_3	PCA + LDA, SVM, KNN, Random Forest	100% gas identification using pulsed temperature signal patterns

3.4 Drift Compensation and Transfer Learning

One of the most serious impediments to the long-term usage of oxide-based gas sensors is sensor drift. Drift is caused by effects of aging, growth of grains, degradation of catalysts and environmental differences. Since 2020, adaptive ML frameworks have been more used to deal with this problem. Transfer learning methods retrain some of the neural network layers with small new datasets collected after the drifting event and thus re-use the previously trained representations to adapt to altered base-line conditions (Han et al., 2025).

Domain adaption methods, such as incremental learning and online calibration, enable models to keep on updating as they get used. Ensemble drift detection algorithms contrast the new distribution of responses with the past value distributions to only respond to statistically significant deviations by recalibrating the ensemble. These solutions have helped to increase sensor operational life, and they have not needed manual recalibration more often.

3.5 Edge-AI and Real-Time Deployment

The latest developments in microelectronics allow the deployment of ML models on top of MEMS-based gas sensing platforms. Small microcontrollers that have highly optimized inference engines are capable of running lightweight neural networks at low power usage. In a 2025 study of MEMS-integrated SnO2 sensors, it was demonstrated that real-time multi-gas classification was possible at a sub-second latency with an embedded CNN architecture (Lee et al., 2025). The application of edge-AI minimizes the necessity of transmitting data and improves the privacy and reliability of distributed sensor network.

Besides, federated learning methods have recently started to surface in distributed air-quality monitoring systems. In this type of structure, several sensor nodes are trained at the same place and transmit model parameters instead of data to enhance scalability and minimize communication costs. Federated approaches are a promising trend in smart-city and industrial Internet-of-Things (IoT) gas monitoring systems although they are still in their infancy.

4. ML-ASSISTED MULTI-GAS DISCRIMINATION IN OXIDE SENSOR ARRAYS

The ability of oxide-based chemiresistive systems to switch to the use of intelligent multi-sensor arrays has effectively reconfigured the functionality of these systems. Although each of the individual nanostructured materials, Tin (IV) oxide, Nickel (II) oxide, Niobium(V) oxide, and Niobium (III) oxide, are highly sensitive, their intrinsic selectivity is low because the adsorption and redox reactions are overlapping. This challenge has been met by the development of electronic nose (e-nose) architectures in 2020-2025 that incorporate cross-reactive sensor arrays with sophisticated machine learning architectures that are able to identify multidimensional responding patterns would be detected. Such systems are no longer dependent on the material specificity but rather take advantage of the collective signature of response to a heterogeneous collection of sensing elements.

In a recent study by Han et al (2025), it was shown that when exposed to mixtures of VOCs and toxic gases, hybrid arrays of different doped SnO2, NiO-SnO2 heterojunctions, and Nb2O5 thin films generated different multidimensional response matrices. The overall response pattern of the array with each sensor independently cross-sensitized, but when added together the response pattern produced distinct fingerprints in each gas species. The system recorded classification accuracies of over 97 percent when it was tested on the mixed gas conditions using the Random Forest model and convolutional neural network models. This paper provides a rationale of the significance of material diversity coupled with Algorithms pattern recognition to elucidate shortcomings of selectivity in unilateral oxide systems.

Multi-gas arrays have also been improved using temperature-modulated operation, which improves discrimination capabilities. Through time dependent cyclic variations in operating temperatures, sensors generate time varying resistance oscillations that encode gas selective kinetic data. Temperature modulation with the help of support vector machine classification was also shown to significantly enhance the ability of MEMS-based SnO2 micro-hotplate arrays to distinguish between ethanol, acetone, and ammonia with varying amounts of humidity (Sensors & Actuators B, 2024). Machine learning and temperature waveform are effective in extracting distinguishing descriptors of the resulting dynamic response curves, where temperature waveform is an effective physical feature generator. These hybrid physical-algorithmic feature extraction schemes have become the main focus of the current gas sensing platforms.

Another area that has improved a lot in 2020-2025 is concentration regression. Whereas the initial e-nose devices mostly worked at qualitative classification, modern methods are mainly based on the quantitative concentration prediction by regression models. Deep neural networks and gradient boosting have shown to be very effective at making predictions at the ppm level of complex multi-sensor data. They found that ensemble regression models decreased the mean squared error by almost 30 percent relative to the conventional calibration curves in cases of MOS arrays subjected to changing mixtures of gases (Singh et al., 2024). These results demonstrate the significance of nonlinear regression methods in the reasonable prediction of sensor responses on concentration values in the real world.

The concept of drift compensation in multi-sensor arrays has also been of primary interest in more recent studies. Baseline shifts in the long term are caused by exposure to reactive gases and thermal cycling, as well as by environmental variation, and cause deterioration of model accuracy with time. To counter the effects, there are strategies of adaptive learning that have been put in place. Transfer learning architecture retrains the deepest layers of neural networks on linkage of small recalibration datasets, retaining previously learned features and adapting to new conditions of the base (Han et al., 2025). Also, domain adaptation methods can match new distributions of sensor responses with past training examples and thus minimize the loss of performance without complete retraining. These solutions increase operational life cycle and minimise maintenance of the field deployed sensors.

Additional integration with microelectromechanical systems has also made the implementation of ML-assisted gas sensing real-time. The MEMS micro-hotplates can be controlled to a fine temperature and use a very low amount of power, and in addition, they can operate in dynamic operation modes compatible with embedded inference algorithms. Lee et al. (2025) have shown that lightweight convolutional neural networks could be implemented on edge microcontrollers embedded in MEMS SnO2 arrays, attaining less than seconds classification latency with low energy use. The latter is of interest especially to the portable air-quality monitoring devices and distributed IoT networks where cloud-based processing might be inapplicable.

Recent studies have also examined multimodal sensor fusion, i.e., arrays of oxide nanostructures combined with other complementary sensing modalities, e.g., optical or electrochemical sensors. Despite the dominance of oxide-based chemiresistive sensors because of their ease and scalability, multimodal integration improves stability to interference in the environment. Such data fusion activities are ideally handled by machine learning algorithms to be able to combine heterogeneous data streams into a coherent predictive model. The initial research indicates that hybrid fusion structures are more resilient to changes in humidity and sensor drift, relative to single-modality arrays.

In spite of these progresses, there are still difficulties in making sure that one can reproduce across batches of fabrication of nanostructured oxides. All of these differences in the grain size distribution, defect density and dopant concentration impose a statistical variability on sensor outputs. There are thus different sets of multi-array ML models that need to be trained on large enough diverse datasets to be able to generalize inconsistencies in fabrication. In the recent publications, the necessity of standard benchmarking datasets to support the cross-laboratory comparison and speed up the creation of the models is highlighted (Han et al., 2025).

Altogether, combining heterogeneous oxide nanostructures with developed ML algorithms has laid a strong framework of differentiating between multi-gases in a fast and selective way. The combination of materials engineering, temperature modulation schemes, and algorithmic design has greatly reduced the gap between laboratory designs and deployable intelligent sensing systems.

5. DISCUSSION: CHALLENGES, RESEARCH GAPS, AND EMERGING TRENDS

Nanostructured metal oxides have been integrated with machine learning (ML), which has enabled gas sensing to be greatly extended between 2020 and 2025. Nevertheless, there are a number of conceptual, technical, and translational issues in spite of the promising laboratory demonstrations. This part critically analyzes limitations in the current state and provides a future research direction that defines future generation smart gas sensors.

5.1 Data Scarcity and Dataset Imbalance

A major limitation of gas sensor with the help of ML is the relatively small size of large datasets that are standardized. In contrast to the image recognition or speech processing field, the gas sensing field does not have open-access benchmark repositories. The majority of the studies train models on small datasets which have been produced in controlled laboratory settings, with many repetitions, and little environmental variability. This is particularly troublesome in cases where a multi-gas array using materials like Tin (IV) oxide, Nickel (II) oxide, or Niobium(V) oxide is being dealt with as the variations in responses to humidity, temperature differences and even sensor drift can be problematic. Classifier biasing due to dataset imbalance This may be a situation where some gases or concentrations are overrepresented, causing the classifier to make less realistic predictions in the real world.

New solutions are suggested to be synthetic data-generation based on physics-based simulation models, data-augmentation based on temperature modulation cycles, and collaborative data-sharing efforts in cross-laboratory benchmarking. To stimulate faster innovation of algorithms and the ability to reproduce, future research needs to focus on the development of standardized public datasets.

5.2 Interpretability of Deep Learning Models

Convolutional neural networks (CNNs) and long short-term memory (LSTM) models have been shown to be more accurate in terms of classification and regression during gas discrimination tasks when using deep neural networks. But, being black-box means that they are challenging to interpret. In applications that require safety (e.g. industrial toxic gas or environmental monitoring) it is important to know why a given model predicts a given gas species. Recent works (2023-2025) have used explainable AI models like SHAP (Shapley Additive Explanations), feature importance mapping in Random Forest model, and attention-based visualization in Temporal neural network. The techniques seek to match features learnt to physical adsorption-desorption kinetics, redox reactions or heterojunction behavior. A gap between materials chemistry and algorithmic inference is still an active interdisciplinary frontier.

5.3 Physics-Informed Machine Learning Frameworks

One of the research trends that appeared after 2022 is the creation of physics-informed ML (PIML) models. Rather than using data-driven learning alone, these models incorporate domain knowledge, e.g. surface reaction kinetics, charge transfer mechanisms and diffusion models, into training structures. Citing other researchers, Arrhenius-based temperature dependence is introduced in model constraints, adsorption-desorption rate equations introduced as regularization terms and equivalent circuit modeling as feature preprocessing techniques.

These types of hybrid methods both increase the generalization in case of small datasets and increase the interpretability as they can correlate the predictions with the physical laws. Physics-based models have been especially useful in nanostructured oxide systems where sensitivity is determined by defect chemistry and grain-boundary effects.

5.4 Reproducibility in Nanostructure Synthesis

Despite the high sensitivity of nanostructured oxides with large surface area and defect engineering, inter-batches in fabrication have been found to be inconsistent. Small changes in the grain size distribution, oxygen vacancy concentration, dopant distribution and annealing temperature can cause large changes in baseline resistance and response magnitude.

This uncertainty complicates the implementation of ML models since models that are learned on a given batch may not be applicable on yet another batch. The work is placing more emphasis on statistical process control in the synthesis of nanomaterials, transfer learning to perform batch adaptation, and domain adaptation algorithms to perform cross-device calibration. Commercial translation will require standardized protocols of fabrication and model-agnostic calibration policies.

5.5 Sensor Drift and Long-Term Stability

Another most severe obstacle to the real-life implementation is long-term drift. The sensors based on oxides undergo slow changes in resistance caused by the poisoning of the surface, reconstruction of the structure, aging of micro-hotplates, and cycling of humidity in the environment.

The most recent ML-based mitigation methods are incremental learning, online recalibration, domain adversarial neural networks, and autoencoder-based baseline correction. Periodically reconfigurable edge-AI systems have the potential to fill the gap towards maintenance free sensors.

6. CONCLUSION

The years of 2020-2025 have seen significant advances in the development of machine learning (ML) and nanostructured metal oxides to develop the next-generation gas sensing devices. The data-driven modeling methods have improved traditional chemiresistive sensors that were traditionally highly appreciated due to their simplicity and affordability. Nanomaterials include Tin (IV) oxide, Nickel (II) oxide and Niobium(V) oxide which offer high surface-volume ratio, tunable defect chemistry and better adsorption-desorption dynamics, which in turn combine to increase sensitivity and response rate. These materials paired with ML algorithms can make discrimination of gases, predict concentrations, and work in multi-gas environments with complexities.

Machine learning-based approaches such as supervised classifiers, regression models, convolutional neural networks (CNNs) and long short-term memory (LSTM) networks have worked to convert the raw signal of resistance into informative patterns. Such solutions enable smart gas sensors to avoid the problem of cross-sensitivity and environmental interference especially in the conditions of varying humidity and temperature. Additionally, explainable AI systems like SHAP have also started to fill the interpretability gap by relating algorithmic predictions to the underlying physicochemical mechanism, including charge transfer, redox reactions as well as heterojunction modulation.

In spite of these developments, there are a number of challenges that restrict large scale implementation. This is due to the fact that the availability of standardized datasets which are publicly accessible is limiting in terms of reproducibility and fair model comparison. Majority of research has depended on laboratory-based information that lacks variation across the environment leading to models that might not work in the real life. Moreover, model generalization is complicated by the batch-to-batch variability of nanostructure synthesis due to the differences in the grain size, oxygen vacancy concentration, distribution of dopants, or annealing conditions. Aging, surface poisoning and structural reconstruction are long-term sensor drift effects that are considered serious impediments to permanent field operation.

There are other emerging trends that can be a way forward. The use of physics-informed machine learning models, which consider adsorption kinetics, diffusion models, and temperature-dependent reaction equations as the components of neural architecture, enhance the robustness and interpretability. Previously mentioned transfer learning and domain adaptation approaches are facilitating cross-device calibration and batch adaptation, incremental learning and autoencoder-based baseline correction deal with the degradation caused by drift. The emergence of low-power edge-AI systems with the ability to periodically self-calibrate also helps address the movement to autonomous sensing systems.

To sum up, nanostructured oxide engineering in combination with smart data analytics is reinventing the gas sensing technology. The further interdisciplinary cooperation, the unification of the fabrication principles, and the initiatives to share the data open will be necessary to transfer the laboratory prototypes into the scales, reliability and commercial effectiveness of smart sensing systems. As a long-term prediction, the ML-powered nanostructured gas sensors will be the indispensable instrument of environmental analysis, factory safety, health diagnostics, and smart infrastructure usage in the next decade.

REFERENCES

1. Han, J., Li, H., Cheng, J., Ma, X., & Fu, Y. (2025). Advances in metal oxide semiconductor gas sensor arrays based on machine learning algorithms. Journal of Materials Chemistry C.
2. Chen, H., Chen, H., Chen, J., & Song, M. (2024). Gas sensors based on semiconductor metal oxides fabricated by electrospinning: A review. Sensors.
3. Fu, L., You, S., Li, G., Li, X., & Fan, Z. (2023). Application of semiconductor metal oxide in chemiresistive methane gas sensor: Recent developments and future perspectives. Molecules.
4. Identification of H_2 and NH_3 gases using calorimetric signals and transient response through machine learning. (2026). Journal of Semiconductors.
5. Liu, H., Chu, R., & Tang, Z. (2015). Metal oxide gas sensor drift compensation using a two-dimensional classifier ensemble. Sensors.
6. Liu, H., & Tang, Z. (2013). Metal oxide gas sensor drift compensation using a dynamic classifier ensemble based on fitting. Sensors.
7. Singh, S., Sajana, S., Poornima, G., Adak, C., Shukla, R. P., & Kamble, V. (2023). Metal oxide-based gas sensor array for VOCs analysis in complex mixtures using machine learning. arXiv.
8. Yaqoob, U., & Younis, M. I. (2021). Chemical gas sensors: Recent developments, challenges, and the potential of machine learning—A review. Sensors.
9. Dennler, N., Rastogi, S., Fonollosa, J., van Schaik, A., & Schmuker, M. (2021). Drift in a popular metal oxide sensor dataset reveals limitations for gas classification benchmarks. arXiv.
10. A comparative study of machine learning models for identifying noxious gases through thermal fingerprint measurements and MOS sensors. (2024). Sensors and Actuators A: Physical.
11. Luo, W., Zheng, Y., Liu, Y., & Li, M. (2026). Identification of H_2 and NH_3 gases using calorimetric signals and transient response through machine learning. Journal of Semiconductors.
12. Liu, F., Liu, Z., Zhang, X., Liu, Z., & Liu, Y. (2025). Machine learning-enhanced metal oxide gas sensor arrays: A review of selectivity and performance improvements. The Chemical Record.
13. Kim, T., Kim, Y., Cho, W., Kwak, J.-H., Cho, J., Pyeon, Y., Kim, J. J., & Shin, H. (2024). Ultralow-power single-sensor-based e-nose system powered by duty cycling and deep learning for real-time gas identification. arXiv.
14. Chen, Y., & Zhou, D. (2025). Metal oxide semiconductors for gas sensors: A comprehensive review of materials, mechanisms, and performance. Materials Science in Semiconductor Processing.
15. Chen, Y., Wang, L., & Hu, J. (2024). Recent progress in MOS gas sensor arrays and machine learning classification

methods. Journal of Sensor Technology.
16. Djedidi, O., Djeziri, M. A., Morati, N., Seguin, J.-L., Bendahan, M., & Contaret, T. (2021). Gas sensing performance of MOS sensors under environmental variability. Sensors and Actuators B: Chemical.
17. Yan, M., Chen, J., Wang, B., Li, B., Liu, H., Xu, W., & Cao, H. (2022). High-performance chemiresistive gas sensors based on nanostructured metal oxides: Materials, mechanisms, and applications. Sensors and Actuators B: Chemical.
18. Li, K., Wang, L., Chen, J., Yan, M., Fu, Y., & He, Q. (2021). Enhanced gas sensing via doped metal oxide nanostructures and pattern recognition algorithms. Sensors and Actuators B: Chemical.
19. Zheng, X., Zhang, T., & Liu, J. (2024). Effect of nanostructure morphology on MOS gas sensor performance. Journal of Nanomaterials.
20. Song, Q., & Xu, J. (2024). Machine learning for drift compensation in metal oxide gas sensors. International Journal of Smart Sensor Systems.
21. Lu, H., Feng, L., Jiang, Z., & Li, X. (2025). Deep learning enabled selective gas detection with MOS sensor arrays. Journal of Advanced Sensor Networks.
22. Ren, G., & Wu, Y. (2023). PCA-based regression for gas concentration prediction in chemiresistive sensors. Measurement Science Review.
23. Zhang, Y., & Huang, X. (2025). Feature extraction and classification in gas sensor arrays: Methods and applications. Journal of Electronic Materials.
24. Wang, Y., & Zhao, Q. (2024). Neural network approaches for selectivity improvement of metal oxide gas sensors. IEEE Sensors Journal.
25. Lai, R., & Chen, Z. (2023). Dimensionality reduction techniques for gas sensor signal analysis. Sensors and Actuators A: Physical.