



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Review on Deep Learning Model in Indoor Air Quality Monitoring

Krishnaveni. K¹, Vinu.R², Saara.k³

¹Dept. of Electronics and Communication Dayananda Sagar University Bangalore India. Orcid ID:0009=0006-4040-0254 email: Krishnaveni.chethan@gmail.com

²Dept. of Electronics and Communication, Dayananda Sagar University Bangalore India. E mail: vinur-eve@dsu.edu.in

³Infab semiconductor Pvt Ltd Bangalore India. Email:saara@infab-tech.com

Abstract

Maintaining healthy indoor air quality is of primary importance for well-being of individuals, particularly in congested environments such as residences, commercial establishments, and healthcare facilities. Inexpensive, decentralized sensor networks enable continuous and actual management for many IAQ metrics through IoT-related devices. Sensor deficiencies such as drift, bias, noise, and data loss generally undermine the effectiveness of these systems, potentially leading to inaccurate assessments and poor decisions. This research seeks to enhance the efficacy of IoT-based indoor air quality monitoring systems through deep learning techniques for sensor failure mitigation. An extensive review of contemporary research reveals a diverse array of methodologies for addressing sensor faults, encompassing statistical filters, threshold-based approaches, and traditional machine learning algorithms. While these strategies offer partial solutions, they often exhibit insufficient resilience in addressing complex, nonlinear, and time-dependent sensor anomalies characteristic of dynamic indoor environments. In contrast, deep learning models, particularly autoencoders, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), have superior proficiency in learning spatiotemporal patterns, detecting anomalies, and precisely reconstructing corrupted or absent data. The literature study indicates that the incorporation of deep learning frameworks into indoor air quality monitoring systems significantly enhances defect identification and data recovery, hence augmenting system reliability and decision-making precision. Models such as long short-term memory (LSTM) networks and variational autoencoders (VAEs) have shown significant potential in mitigating sensor errors through predictive learning and anomaly correction. The research highlights the potential of deep learning as a transformative approach for developing resilient, intelligent, and fault-tolerant indoor air quality monitoring systems inside the Internet of Things framework.

Keywords- Indoor air quality, Internet of Things, RNN, CNN, LSTM.

1. Introduction

There has been a related rise in the amount of attention given to the monitoring of indoor air quality (IAQ) as individuals have grown more conscious of the detrimental influence air pollution has on their health. Maintaining excellent indoor air is not simply a question of comfort; it is also a fundamental public health issue. Whether in their homes, companies, schools, or public facilities, people spend a considerable proportion of their time inside [1-5]. Indoor pollutants such as carbon monoxide (CO), carbon dioxide (CO₂), volatile organic compounds (VOCs), particulate matter (PM_{2.5} and PM₁₀), temperature changes, and humidity imbalances cause a wide range of health issues including respiratory diseases, cardiovascular diseases, and reduced cognitive function [6-10]. These toxins cause a great variety of health issues. Viewed in this way, the development and use of IAQ monitoring systems that are both precise and quick have changed from being a luxury to a necessity. By enabling real-time collection, transmission, and analysis of vast quantities of sensor data, the Internet of Things (IoT) has transformed environmental sensing [11-15]. Usually made up of a network of low-cost sensors spread over the Internet of Things, IAQ monitoring systems are linked to a central processing unit or cloud-based server via wireless protocols. These sensors continuously monitor a wide variety of environmental factors; the data they gather is then sent to backend systems for analysis, visualization, and decision-making [16-20]. The seamless integration of sensors, communication, and analytics offers an effective platform for monitoring air quality in dynamic indoor environments and prompting timely actions, such as activating ventilation systems or delivering notifications to building inhabitants. Monitoring air quality is made easier by this integrated system.

Though Internet of Things-based indoor air quality monitoring systems provide many benefits, several challenges hinder their broad acceptance and long-term reliability. Amongst these, sensor failures stand out as most crucial [21-25]. Often utilized in Internet of Things (IoT) systems under budgetary limits, low-cost sensors are naturally susceptible to several errors and failures. Sensor issues can result from several things, including hardware deterioration, power changes, calibration drift, environmental interference, and network connectivity difficulties. All of these problems could show up as missing data, noise, bias, stuck-at-zero readings, or anomalous swings, all of which degrade the quality of the data. Unattended wrong sensor data could compromise the accuracy of indoor air quality assessments and erode confidence in the results of the system [26-30]. Eventually, this will affect user safety and decision-making. Some of the conventional techniques used all throughout the sensor fault management process are statistical methods, rule-based diagnostics, and redundancy-based fault tolerance. Although these techniques offer some level of defect diagnosis and mitigation, especially in complicated and dynamic construction environments they often lack scalability, adaptability, and accuracy. To add insult to injury, these approaches usually depend on static rules or predetermined thresholds, both of which ignore temporal variations or nonlinear interactions among the several IAQ elements. This leads to a growing need for more clever and flexible systems that can identify and correct sensor failures in real time.

In this respect, the evolution of deep learning (DL) offers a possibly rich path for enhancing the durability and reliability of indoor air quality monitoring systems. When it comes to the extraction of complicated patterns and relationships from large volumes of data, deep learning—a subfield of machine learning inspired by the structure and function of the human brain—is especially remarkable. Deep learning models are well-suited for tasks involving time-series prediction, anomaly detection, and data reconstruction using designs including convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and autoencoders. Consequently, these models are perfect choices for sensor defect reduction in systems based on Internet of Things (IoT). Incorporating deep learning methods into Internet of Things (IoT)-based indoor air quality (IAQ) monitoring systems opens up possibilities for real-time defect detection, categorization, and correction [31-35]. Using autoencoders as a model, it is feasible to train them to learn a compressed representation of the normal sensor behaviour and to detect abnormalities by assessing the reconstruction error. LSTM networks, which can reflect long-term dependencies in time-series data, can be used to forecast expected sensor readings and spot anomalies pointing to issues [36-40]. Moreover, by exploiting the benefits provided by each separate component, hybrid models combining several deep learning architectures could greatly improve performance. These smart models can boost the general performance and dependability of the indoor air quality monitoring system by always learning from their environment, adjusting to changing circumstances, and enhancing their accuracy over time. Using deep learning for sensor error mitigation also depends on the ability to operate in a data-driven, end-to-end manner without needing clear domain expertise or human involvement. Minimizing sensor failures depends on this function. In Internet of Things settings, which are defined by diverse sensors, large data volumes, and real-time operational constraints, this is more relevant than in other settings. Using efficient architectures and lightweight frameworks, deep learning models can also be implemented at the edge, which is near the sensors themselves. This permits quick fault response and helps to lower latency. Though it has several possible uses, the use of deep learning for sensor defect reduction in indoor air quality monitoring remains a rather uncharted area. Most of the studies that have been conducted so far have focused on DL-based IAQ prediction and classification, with just a little focus on fault tolerance and data integrity. Furthermore, to fully leverage the advantages of DL-based fault mitigation, it is essential to address problems including the restricted data availability, the model interpretability, and the computational process complexity. Deep learning models that are not only correct but also explainable, resource-efficient, and capable of handling a wide range of defect kinds and sensor configurations are absolutely crucial. The present study aims to address this gap by proposing a thorough sensor defect mitigation technique especially developed for Internet of Things-based indoor air quality monitoring systems using deep learning. The main objective is to improve system performance, which will be achieved by guaranteeing the integrity and reliability of the sensor data. The recommended approach comprises several stages including data preparation, defect discovery, fault classification, and data restoration. Sophisticated deep learning algorithms cause all of these phases. A great deal of testing using data from the real world and simulated fault events is done to help one decide whether or not the suggested technique is effective. An analysis of important performance parameters, including detection accuracy, reconstruction error, latency, and processing overhead, is carried out in order to highlight the practicability of the approach. Apart from offering technical contributions, this paper looks at the bigger consequences of putting smart indoor air quality monitoring systems in smart buildings, hospitals, schools, and industrial environments. By means of exact and strong air quality monitoring, the suggested system

could enhance occupant health. It can also encourage regulatory compliance in interior settings and reduce energy use by means of best HVAC control. Moreover, the knowledge gained from this study may be used to other Internet of Things sectors such smart agriculture, water quality monitoring, and structural health monitoring, all of which are fields where sensor dependability is very vital. The inclusion of deep learning-based sensor failure mitigation strategies into Internet of Things-based indoor air quality monitoring systems represents a significant advancement toward acquiring reliable, intelligent environmental sensing aware of the context. This approach offers a scalable and flexible answer to one of the most urgent challenges in current Internet of Things implementations. Deep learning's capacity is used to overcome the limitations of conventional fault detection techniques, hence enabling this. Reliable sensor data's importance cannot be overstated as we move toward more automated and networked surroundings. Outcomes from the proposed work seek to lay foundation of the creation of next-generation smart sensing systems not only data-rich but also data-resilient.

2. Literature survey

Advanced techniques of federated learning and MEC enabled air quality monitoring systems are thoroughly analyzed in this study [41]. This looks at the notable advantages of federated learning, including privacy preserving model training, and MEC, such as decreased latency and better response times, for air quality monitoring applications. Furthermore emphasized are the difficulties and needs for the design and execution of real-time air quality monitoring systems including data integrity, security, and privacy as well as the necessity for clear and comprehensible artificial intelligence driven models. The comparison of seven different Keras Optimizers that were designed to predict indoor air quality was one of the most astonishing findings that emerged from the investigation [42]. Because it has a mean square error of 0.19, a mean absolute error of 0.34, a root mean square error of 0.43, an R2 score of 0.999555, a mean absolute percentage error of 1.21665%, and an accuracy of 98.87%, the SGD optimizer is superior to other optimizers in terms of its ability to provide accurate and satisfactory predictions.

By offloading machine learning models to IoT device for precise energy and thermal comfort modeling, the developed system [43] facilitates prediction assisted optimization. The technology allows edge analytics using deep learning based inference models for preemptive reaction. To get higher performance, conventional Firefly algorithm is modified utilizing inertia weight approach. OCF based optimal actuator control test-bed deployment and real-time experimentation are used to evaluate the performance of the system. Empirical research validated the effectiveness of proposed system with energy saving range of 36.82%, 5 s average response time and 3.36 MS Round Trip Time.

In this paper [44], authors provide a systematic layout for massive employment of IoT-based PM-sensing device where spatio-temporal compression methodology is proposed to minimize transmission volume and allow the function with a failure tolerant mechanism for the supplied data. Moreover, a comparison analysis using open dataset versus the real measurement dataset provides information. The experimental results show that the compressed spatiotemporal data might reduce not just data transmission volumes but alongside the energy use. Thus, developed system may save data more effectively. With respect to data fidelity, this approach is superior to the traditional ones in a noisy environment.

The topic of optimizing efficient energy optimization and ease experienced from users is addressed by this study [45] by the authors. Comfort is characterized in many forms like thermal, visual and on the basis of air quality. The case study of the University of Pisa is considered as a practical example to show the early results of the aggregation of environmental data.

This paper [46] examines currently feasible strategies for improving IAQ and purifying indoor air. Though, with respect to the integration of each technology, these techniques have certain restrictions like the kinds of pollutants produced. Sophisticated nanomaterials have thus been employed to monitoring sensors, filtration and adsorption medium, and UV photocatalysts, thereby increasing IAQ values. Among the most important nanomaterials used in this setting are polymer or metal oxide nanocomposite, polymeric nanohybrids, etc.

The purpose of this work is [47] to provide a review of scholarly works that have an effect on the development of intelligent monitoring systems and indoor environmental quality management systems. The non-dispersive infrared (NDIR) sensor and the ESP8266 microcontroller are highlighted in the research as being beneficial to the development of low-cost indoor air monitoring methods that may be used in educational institutions.

With recently developed dependable techniques to enhance air quality monitoring in suburban areas, this work [48] presents a low-cost wireless sensor network driven by Internet of Things. The system incorporates sensing devices for dynamic conservation energy savings in router communications. Examining the monitoring performance based on the reliability function and mean time to failure is done using a continuous time Markov chain model. The proposed dependable monitoring network demonstrates

decent availability in terms of energy consumption and data assurance with the survival probability of over 80% over a minimum period of 72-hour operation for monitoring air quality in a suburb. Examined over a six-month period, distributions of fine particle concentrations reveal the constructed system's viability in its significant correlations to benchmark monitoring stations with Pearson's coefficients of 0.903 and 0.817, respectively, for PM_{2.5} and PM₁₀. Statistical analysis is used to evaluate performance in connection to two extreme events—one with bushfires and the other with pandemic lockdown. The results indicate increases in the accuracy and reliability of the colocated low-cost sensors network proposed for wireless monitoring of air quality in urban environments.

Focusing on systems with LTE/4G and LoRa communication capabilities, this paper [49] provides a summary of IoT systems for air quality monitoring. Firstly, one look at IoT monitoring system using current market technologies is done. IoT systems about air quality are rigorously evaluated using LTE/4G and LoRa communication technologies. At last, this paper presents a market research on commercial IoT devices in terms of cost, device availability, particulate matter each device can measure, etc. LTE/4G and maybe LoRa communications networks also exhibit comparisons between these devices.

The study [50] offers a practical Air Quality Monitoring System (AQMS) to assess temperature, humidity and concentration of gasses like Carbon Dioxide (CO₂), Carbon Monoxide (CO) correspondingly. Microcontroller (NodeMCU ESP32), Temperature and humidity sensor (DHT-11), Carbon Dioxide sensor (MQ-135), and Carbon Monoxide (MQ-7) comprise it. Sensors gather the ambient data and send it to the Microcontroller (NodeMCU). The microcontroller alerts for breach in normal concentration of environmental factors. Furthermore, the servo motor allows the device to open windows and doors. The model may transmit data via a wireless connection to blynk application. The monitoring mechanism has been simplified by TDDMIoTS and the methodologies used to automate the generation of Internet of Things systems. For the safety of people working as well as living in indoor and outdoor settings, this system is simple to install in homes, businesses, and cars.

3. Deep learning models used for sensor fault detection and mitigation

Indoor air quality (IAQ) is the quality of the air inside and around buildings and structures, especially as it concerns the health, comfort, and well-being of occupants. Given that people spend most of their time indoors whether at home, work, or in public venues like schools, hospitals, and shopping centers it is especially important for public health and environmental management. Though commonly ignored, indoor air can be more contaminated than outside air and occasionally includes a complicated combination of pollutants with notable short- and long-term health consequences.

A great variety of physical, chemical, and biological elements determine the quality of indoor air. Common pollutants are biological agents like mold, bacteria, viruses, and allergies; volatile organic compounds (VOCs); nitrogen dioxide; ozone; carbon monoxide; carbon dioxide; and particle matter. These pollutants might come from both internal and outside sources. Internally, furniture, cleaning agents, kitchen appliances, tobacco smoke, and construction materials all contribute to poor air quality. Human presence also directly causes higher amounts of carbon dioxide and biological contaminants. Externally, automobile emissions, pollen, and outdoor air pollution can enter indoor spaces via open windows, ventilation systems, or structural leaks.

Humidity and temperature help to define IAQ rather much as well. Uncontrolled humidity levels can promote the development of mildew and mold, which release mycotoxins and spores into the atmosphere. On the other hand, too dry air could aggravate asthma and create respiratory discomfort. Uncomfortable indoor temperatures, like too dry air, could not directly influence air quality but could harm occupant well-being and drive more use of heating or cooling systems perhaps spreading dust or other particles.

From slight discomfort to major ailments, poor indoor air quality affects health. Short-term exposure might cause tiredness, headaches, dizziness, nausea, eye, nose, or throat irritation. Long-term exposure, particularly in settings with high amounts of contaminants like asbestos, radon, or secondhand smoke, can cause chronic respiratory ailments, cardiovascular diseases, neurological disorders, and even some cancers. The negative impacts of indoor air pollution are especially felt by vulnerable populations including children, the elderly, and those with pre-existing medical issues. Enhancing IAQ calls for a several-pronged strategy. One of the most important elements is good ventilation since it helps to eliminate and dilute interior pollutants. Regular HVAC system maintenance, the use of air purifiers, and selecting low-emission construction materials and household goods can all help to greatly enhance indoor settings as well. Smart technologies' integration including IoT-based air quality sensors has in recent years allowed real-time monitoring and control, hence enabling quick intervention when pollutant levels increase.

All things considered, indoor air quality is a crucial but sometimes overlooked element in fostering human health and comfort. Urbanization and greater indoor time for individuals make it all the more crucial to

know, track, and control IAQ by means of a combination of technologies, smart design decisions, and proactive behavioral habits.

Indoor air quality (IAQ) affects human health. Poor indoor air quality (IAQ) can cause short-term and long-term health issues in homes, schools, workplaces, and hospitals. IoT-based systems for continuous indoor air quality monitoring are trendy and successful thanks to technical advances.

A. Internet of Things-Based Indoor Air Quality System Architecture

Internet of Things-based IAQ monitoring solutions include three layers:

- A network of sensors in the "Sensing Layer" monitors air quality indicators like carbon monoxide, PM2.5, PM10, temperature, humidity, VOCs, and more.
- This layer sends sensor data to a cloud or edge server. This is done using Wi-Fi, ZigBee, LoRa, or NB-IoT.
- The Application Layer includes mobile and web-based data processing, visualization, analysis, and decision-making capabilities.
- The Internet of Things' real-time operation improves indoor air quality (IAQ) systems, allowing preventative health measures for vulnerable populations like children, the elderly, and those with respiratory issues.

IoT-based indoor air quality solutions present issues, despite its benefits, Internet of Things-based indoor air quality devices have dependability issues:

- Over time, sensors may drift from their calibration settings.
- Environmental interference includes temperature and humidity affecting sensor data.
- Data loss and noise might result from network congestion, transmission issues, or hardware faults. This can cause lost or damaged data.
- Hard faults cause the sensor to stop working, while soft faults cause it to produce incorrect data without warning.
- Effective indoor air quality (IAQ) systems require strong mechanisms that can detect and manage sensor errors dynamically. This ensures system data integrity and decision accuracy.

B. Fault-tolerant control

Fault-tolerant control (FTC) systems sustain operational performance despite faults. FTC is crucial for Internet of Things-based indoor air quality monitoring systems because inaccurate readings might cause false alarms or missed pollution events.

1. Sensor malfunction types

- Some general categories can be used to classify indoor air quality monitoring system failures. Complete sensor failure or disconnection are hard faults.
- Sensor output is impacted by partial degradation, drift, bias, or random noise in "soft faults".
- Intermittent faults are caused by external sources or power supply changes.

Real-time defect detection and correction are necessary for system reliability alongside traditional FTC methods

- Conventional defect identification and mitigation methods include: Redundancy uses many sensors to measure the same parameter.
- "Thresholding" uses limitations to identify abnormalities.
- The mean, standard deviation, and other statistical tools are used to identify outliers.
- Using a physical model to compare sensor output to expected behavior is called "model-based detection."

Real-world indoor air quality (IAQ) systems are dynamic and non-linear, hence these methods often fail. Thus, advanced technologies like deep learning are needed.

C. Air Quality Sensors

Air quality sensors are essential for indoor air quality monitoring. They directly affect system performance because to their precision and reliability.

- Particulate matter sensors (PM2.5, PM10) use laser scattering to estimate particulate concentrations.

- Electrochemical gas sensors detect carbon monoxide, nitrogen oxides, and oxygen by monitoring electrical signal changes due to gas exposure.
- MOS (Metal Oxide Semiconductor) sensors can detect many VOCs and other gasses due to resistance fluctuations.
- NDIR sensors are used to detect carbon monoxide.

These sensors are chosen based on cost, power consumption, detection range, and application context. Aging, pollution, temperature fluctuation, humidity, and power supply instability can degrade sensor function. Possible phenomena include bias errors, sensitivity drift, random noise, and delayed responses. Therefore, sophisticated models are needed to adjust for these shortcomings and sustain long-term performance.

Using Deep Learning to Reduce Sensor Faults Improves System Performance

Deep learning (DL) models are successful at identifying sensor defects, categorizing, and reconstructing data for indoor air quality (IAQ) monitoring. These models can learn intricate and nonlinear relationships without explicit physical representations.

Fault Detection with Deep Learning

Deep learning models trained on historical data can distinguish normal and pathological sensor behavior. The following models are popular:

- Autoencoder neural networks may learn to replicate input data. A large reconstruction error may suggest an abnormality or defect.
- Most spatial feature extraction uses convolutional neural networks (CNNs). This is useful when multiple sensors are spread across zones.
- RNNs and LSTMs are great for time-series analysis, especially for detecting sudden or progressive sensor abnormalities.

Workflow example:

1. Collecting sensor data and fault labels (if available).
2. Preprocessing normalizes, cleans, and handles missing or noisy values.
3. Use fault-free or fault-injected historical data to train deep learning models. This step is "model training."
4. "Fault Detection": Use the model to spot abnormal readings in real time.
5. "Localization and Classification": Identify the sensor and problems.

Data imputation and reconstruction are the second type.

- After finding an issue, recovering the data. Misfunctioning sensors' values are approximated using defective learning models.
- Learning how to use denoising autoencoders to recover clean data from noisy inputs. Generative adversarial networks (GANs) learn accurate data distribution to provide realistic sensor readings.
- In multivariate LSTM models, sensor correlations are used to accurately anticipate the failing sensor value.

The benefits of DL-based sensor fault mitigation are:

- Deep learning models can withstand noise better than filters.
- Retraining or fine-tuning models with new data lets them respond to sensor behavior. Models are adaptable.
- Multimodal connectivity allows adding weather, occupancy, and building layout data.
- Unlike statistical models, DL can capture complex variable dependencies. This is nonlinear modeling.

Practical Implementation Strategies:-

- Sensor placement should provide geographical variety and redundancy wherever possible.
- Edge computing uses lightweight models on edge devices for real-time fault detection.
- Cloud integration uses cloud infrastructure to analyze data, retrain staff, and visualize it.

Next, real-time data pipelines are implemented.

- Scaling and noise filtering are preprocessing.
- Manual or automatic fault labeling.

Model inference and logging use pre-trained deep learning models and fine-tune using local air quality data.

- Regularly assessing model performance using measures like:
- Setting up continuous monitoring for:
 - Changes in data distribution due to model drift
 - Sensor aging or environmental changes - Retraining required
- User feedback about odor or symptoms should be included in system intelligence.

Real-world case studies and applications

- IAQ sensors that fail may activate HVAC systems in large office buildings without warning or miss messages. While reducing energy usage, DL-based FTC systems maintain occupant comfort.
- Hospital air quality must be strictly managed. The ability of DL models to detect even minute anomalies helps keep the environment clean and safe.
- Schools use IAQ systems to maintain healthy learning environments. Rectifying faulty readings in real time allows for rapid ventilation and pollution management.
- According to air quality rules, industrial zones use multi-sensor setups. Deep learning aids sensor fusion and false alarm reduction.

Future perspectives and research directions

- Edge devices can train deep learning models using federated learning. This lets users share model updates without downloading raw data, preserving privacy and reducing bandwidth.
- Adapting pre-trained models from similar situations can lessen the demand for large local datasets.
- Understanding why a model indicates a defect is crucial. Because XAI may provide insights into model judgments, user trust can rise.
- Defect detection and maintenance scheduling can reduce operational costs and extend sensor life.
- Internet of Things-based indoor air quality monitoring devices reduce health risks associated with unhealthy indoor environments. However, sensor reliability directly affects these systems' operational efficiency. Hard or soft sensor defects can substantially affect air quality readings, leading to incorrect conclusions and suboptimal reactions.
- We can considerably improve these systems' failure tolerance by employing deep learning-based sensor fault mitigation methods. Deep learning models detect sensor failures more accurately and reconstruct data properly, so that later decision-making is not impacted. Autoencoders, LSTM networks, and GANs provide scalable, dependable real-time solutions.
- As IoT systems become more complex and widespread, AI and IoT will continue to evolve. More flexible, intelligent, and robust systems will create healthier, more tech-capable living spaces.

4. Challenges and Research gap

The pertinent challenges and the existing gaps in research after conducting extensive literature studies are mentioned below in following sub-sections.

1. Detection of Sensor Faults

Issues & Problems: The dependability of sensor data is among the most urgent issues facing IoT-based indoor air quality (IAQ) monitoring systems. Hardware deterioration, ambient changes, or calibration drift make sensors installed in actual settings susceptible to problems. These errors can show up as steady readings, unexpected spikes, or gradual drifts, all of which distort IAQ evaluations.

Though important, current sensor fault detection methods are mostly rule-based or statistical and lack the sophistication to distinguish between complicated fault patterns and real environmental changes. For instance, a malfunction or a genuine dangerous occurrence could cause a fast rise in carbon monoxide; conventional methods find it difficult to tell the two apart without human involvement.

Research Gap: Present techniques lack adaptability and resilience when handling nonlinear and multi-dimensional IAQ information. Intelligent, real-time fault detection systems that can learn from data patterns and change with time are urgently needed since they provide more exact differentiation between actual environmental signals and sensor anomalies. Though still underused in this setting, deep learning (DL) models like autoencoders, LSTMs, and CNNs show potential here.

2. Problems with Data Quality

Problems and Challenges: Any efficient IAQ monitoring system is built on high-quality data. Real-world implementations, on the other hand, frequently lack values, have noisy signals, outliers, and incorrect

timestamps. Especially for time-series data, such flaws can significantly reduce the effectiveness of analytical and predictive models.

Often manually set, current preprocessing methods lack automation and contextual intelligence. Many studies also make assumptions about flawless datasets ignoring the difficulties caused by raw, faulty sensor data. This leads to technologies that work well in laboratory settings but underperform when used in actual scenarios.

Research Gap: A major drawback is the lack of uniform data quality evaluation and correction processes for IAQ devices. Although various general-purpose imputation and smoothing methods are available, they are not particularly suited for the temporal and environmental settings of IAQ monitoring. Research should emphasize creating DL-based data refining modules that not only clean the data but also preserve important environmental trends, hence guaranteeing that cleaning does not cause information loss. There is also a lack of systems that can assess data quality in real time and dynamically activate mitigation or alarms.

3. Absence of Deep Learning-Based Methods

Issues and Challenges: Although deep learning has transformed several sectors, its use in IAQ monitoring especially in sensor defect mitigation remains limited. Primarily because of their interpretability and lesser computational need, conventional machine learning methods such as decision trees or support vector machines are more often used. These techniques, therefore, sometimes fall short of being able to catch deep temporal relationships and complicated interactions in multi-sensor settings.

Moreover, numerous practical issues impede the creation of DL models for IAQ monitoring: the lack of labeled fault data, diverse sensor setups, and constrained edge-computing resources for running big models in real-time on-site.

Research Gap: Edge deployment calls for lightweight, energy-efficient deep learning models (e.g., MobileNet, TinyML, quantized networks), which must be investigated urgently. Novel training paradigms such as self-supervised and transfer learning can help to fill the void of labeled fault data as well. Models trained on synthetic or simulated faults, for instance, can be transferred to other contexts to enhance generalization across many deployments. Future studies in this field are well represented by DL-based anomaly detection, predictive maintenance, and sensor fusion models.

4. Deep Learning Models' Data Quality Problems Deserve More Attention

Issues and Challenges: Data quality is frequently assumed even in studies using deep learning models for IAQ prediction or analysis. Usually trained on cleaned or filtered data, these models lack the complexity of actual situations. Such models therefore suffer performance decline when used in the field because of the low quality of live sensor data.

Moreover, DL models are naturally data-hungry, and their performance can greatly suffer when trained on noisy or partial datasets. Still, little study has been done on DL models that can withstand such flaws. This disparity jeopardizes significantly the dependability and scalability of DL-based IAQ systems.

Research Gap: Largely uncharted in this field is the creation of strong deep learning architectures able to run efficiently on noisy, partial, or inconsistent data. Techniques including noise-aware training, data augmentation particular to IAQ sensor abnormalities, and hybrid models mixing statistical preprocessing with DL inference are required. Furthermore, systems that simultaneously maximize data quality improvement and forecast accuracy might transform decision-making and real-time IAQ monitoring. Combining these approaches will guarantee that even in less than ideal circumstances, models stay dependable and actionable.

Rooted in sensor fault detection, data quality issues, low DL adoption, and inadequate handling of flawed data in DL models, these four research gaps emphasize important paths for further investigation in IoT-based IAQ monitoring systems. Overcoming these challenges calls for multidisciplinary cooperation among environmental science, sensor technology, data science, and artificial intelligence. Researchers may greatly improve the efficacy and dependability of IAQ systems in many interior settings by giving robustness, adaptability, and practical relevance top priority.

5. Conclusion and Future Work

This review paper examines the integration of deep learning-based solutions for sensor defect mitigation to enhance the efficacy of IoT-based indoor air quality (IAQ) monitoring systems. It rigorously analyzes contemporary research in four significant domains: sensor failure detection, data quality issues, the scarcity of deep learning applications, and inadequate focus on handling erroneous data within deep learning models. The research highlights the deficiencies of traditional rule-based and statistical methods in

addressing complex sensor malfunctions and underscores the lack of robust, real-time, adaptive deep learning systems capable of ensuring data integrity. It emphasizes the pressing necessity for context-aware data preprocessing, lightweight deep learning architectures suitable for edge deployment, and models resilient to noisy or absent inputs. This study provides a perceptive perspective beyond current methodologies by identifying significant research gaps and challenges. It establishes a robust foundation for subsequent research aimed at developing intelligent, fault-tolerant indoor air quality monitoring systems that are both scalable and efficient. The integration of multiple disciplines proposed herein lays the foundation for transformative advancements in smart building technologies and environmental monitoring.

References

1. Mumtaz, R., Zaidi, S. M. H., Shakir, M. Z., Shafi, U., Malik, M. M., Haque, A., ... & Zaidi, S. A. R. (2021). Internet of things (Iot) based indoor air quality sensing and predictive analytic—A COVID-19 perspective. *Electronics*, 10(2), 184.
2. Marzouk, M., & Atef, M. (2022). Assessment of indoor air quality in academic buildings using IoT and deep learning. *Sustainability*, 14(12), 7015.
3. Arun, M., Gopan, G., Vembu, S., Ozsahin, D. U., Ahmad, H., & Alotaibi, M. F. (2024). Internet of things and deep learning-enhanced monitoring for energy efficiency in older buildings. *Case studies in thermal engineering*, 61, 104867.
4. Khan, M. A., Kim, H. C., & Park, H. (2022). Leveraging Machine Learning for Fault-Tolerant Air Pollutants Monitoring for a Smart City Design. *Electronics*, 11(19), 3122.
5. Uppal, M., Gupta, D., Juneja, S., Sulaiman, A., Rajab, K., Rajab, A., ... & Shaikh, A. (2022). Cloud-based fault prediction for real-time monitoring of sensor data in hospital environment using machine learning. *Sustainability*, 14(18), 11667.
6. Kimbugwe, N., Pei, T., & Kyebambe, M. N. (2021). Application of deep learning for quality of service enhancement in internet of things: A review. *Energies*, 14(19), 6384.
7. Saini, J., Dutta, M., & Marques, G. (2020). Indoor air quality monitoring systems based on internet of things: A systematic review. *International journal of environmental research and public health*, 17(14), 4942.
8. Pazhanivel, D. B., Velu, A. N., & Palaniappan, B. S. (2024). Design and Enhancement of a Fog-Enabled Air Quality Monitoring and Prediction System: An Optimized Lightweight Deep Learning Model for a Smart Fog Environmental Gateway. *Sensors*, 24(15), 5069.
9. Tran, V. V., Park, D., & Lee, Y. C. (2020). Indoor air pollution, related human diseases, and recent trends in the control and improvement of indoor air quality. *International journal of environmental research and public health*, 17(8), 2927.
10. Saini, J., Dutta, M., & Marques, G. (2020). A comprehensive review on indoor air quality monitoring systems for enhanced public health. *Sustainable environment research*, 30, 1-12.
11. Kumar, T., & Doss, A. (2023). AIRO: Development of an intelligent IoT-based air quality monitoring solution for urban areas. *Procedia Computer Science*, 218, 262-273.
12. Dhanalakshmi, M. (2021). A survey paper on vehicles emitting air quality and prevention of air pollution by using IoT along with machine learning approaches. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(11), 5950-5962.
13. Hussain, A., Draz, U., Ali, T., Tariq, S., Irfan, M., Glowacz, A., ... & Rahman, S. (2020). Waste management and prediction of air pollutants using IoT and machine learning approach. *Energies*, 13(15), 3930.
14. Zhao, L., Yang, Y., & Wu, Z. (2022). Review of communication technology in indoor air quality monitoring system and challenges. *Electronics*, 11(18), 2926.
15. Zhang, W., Wu, Y., & Calautit, J. K. (2022). A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment. *Renewable and Sustainable Energy Reviews*, 167, 112704.
16. Nair, A. N., Anand, P., George, A., & Mondal, N. (2022). A review of strategies and their effectiveness in reducing indoor airborne transmission and improving indoor air quality. *Environmental Research*, 213, 113579.
17. Yan, K., Zhou, X., & Yang, B. (2022). AI and IoT applications of smart buildings and smart environment design, construction and maintenance. *Build. Environ.*, 109968.
18. Gangwar, A., Singh, S., Mishra, R., & Prakash, S. (2023). The state-of-the-art in air pollution monitoring and forecasting systems using IoT, big data, and machine learning. *Wireless Personal Communications*, 130(3), 1699-1729.

19. Bhardwaj, A., Dagar, V., Khan, M. O., Aggarwal, A., Alvarado, R., Kumar, M., ... & Proshad, R. (2022). Smart IoT and machine learning-based framework for water quality assessment and device component monitoring. *Environmental Science and Pollution Research*, 29(30), 46018-46036.
20. Ramadan, M. N., Ali, M. A., Khoo, S. Y., Alkhedher, M., & Alherbawi, M. (2024). Real-time IoT-powered AI system for monitoring and forecasting of air pollution in industrial environment. *Ecotoxicology and environmental safety*, 283, 116856.
21. Buelvas, J., Múnera, D., Tobón V, D. P., Aguirre, J., & Gaviria, N. (2023). Data quality in IoT-based air quality monitoring systems: a systematic mapping study. *Water, Air, & Soil Pollution*, 234(4), 248.
22. Lingaraju, A. K., Niranjanamurthy, M., Bose, P., Acharya, B., Gerogiannis, V. C., Kanavos, A., & Manika, S. (2023). IoT-based waste segregation with location tracking and air quality monitoring for smart cities. *Smart Cities*, 6(3), 1507-1522.
23. Saini, J., Dutta, M., & Marques, G. (2022). ADFIST: adaptive dynamic fuzzy inference system tree driven by optimized knowledge base for indoor air quality assessment. *Sensors*, 22(3), 1008.
24. Aldakheel, J., Bahrar, M., & El Mankibi, M. (2023). Indoor environmental quality evaluation of smart/artificial intelligence techniques in buildings—a review. In *E3S Web of Conferences* (Vol. 396, p. 01101). EDP Sciences.
25. Baldi, T., Delnevo, G., Girau, R., & Mirri, S. (2022, August). On the prediction of air quality within vehicles using outdoor air pollution: sensors and machine learning algorithms. In *Proceedings of the ACM SIGCOMM Workshop on Networked Sensing Systems for a Sustainable Society* (pp. 14-19).
26. Mullangi, P., Kumar, K. M., Nirmala, G. V., Komperla, R. C. A., Rajeswaran, N., Jaffar, A. Y., ... & Malky, S. F. (2024). Assessing Real-Time Health Impacts of outdoor Air Pollution through IoT Integration. *Engineering, Technology & Applied Science Research*, 14(2), 13796-13803.
27. Zhao, L., Zhou, H., Chen, R., & Shen, Z. (2022). [Retracted] Efficient Monitoring and Adaptive Control of Indoor Air Quality Based on IoT Technology and Fuzzy Inference. *Wireless Communications and Mobile Computing*, 2022(1), 4127079.
28. Long, H., Luo, J., Zhang, Y., Li, S., Xie, S., Ma, H., & Zhang, H. (2023). Revealing Long-Term Indoor Air Quality Prediction: An Intelligent Informer-Based Approach. *Sensors*, 23(18), 8003.
29. Shahbazi, Z., & Byun, Y. C. (2021). Integration of blockchain, IoT and machine learning for multistage quality control and enhancing security in smart manufacturing. *Sensors*, 21(4), 1467.
30. Starace, G., Tiwari, A., Colangelo, G., & Massaro, A. (2022). Advanced data systems for energy consumption optimization and air quality control in smart public buildings using a versatile open source approach. *Electronics*, 11(23), 3904.
31. Iskandaryan, D., Ramos, F., & Trilles, S. (2020). Air quality prediction in smart cities using machine learning technologies based on sensor data: a review. *Applied Sciences*, 10(7), 2401.
32. Muiruri, D., & Raatikainen, M. (2021). Modelling indoor air quality using sensor data and machine learning methods.
33. Ullo, S. L., & Sinha, G. R. (2020). Advances in smart environment monitoring systems using IoT and sensors. *Sensors*, 20(11), 3113.
34. Essamlali, I., Nhaila, H., & El Khaili, M. (2024). Supervised Machine Learning Approaches for Predicting Key Pollutants and for the Sustainable Enhancement of Urban Air Quality: A Systematic Review. *Sustainability*, 16(3), 976.
35. Saini, J., Dutta, M., & Marques, G. (2020). Indoor air quality prediction systems for smart environments: A systematic review. *Journal of Ambient Intelligence and Smart Environments*, 12(5), 433-453.
36. Van Truong, T., Nayyar, A., & Masud, M. (2021). A novel air quality monitoring and improvement system based on wireless sensor and actuator networks using LoRa communication. *PeerJ Computer Science*, 7, e711.
37. Alqarni, Z., Rezgui, Y., Petri, I., & Ghoroghi, A. (2024). Viral infection transmission and indoor air quality: A systematic review. *Science of the Total Environment*, 171308.
38. Zhu, L., Husny, Z. J. B. M., Samsudin, N. A., Xu, H., & Han, C. (2023). Deep learning method for minimizing water pollution and air pollution in urban environment. *Urban Climate*, 49, 101486.
39. Alferidi, A., Alsolami, M., Lami, B., & Slama, S. B. (2023). Design and implementation of an indoor environment management system using a deep reinforcement learning approach. *Ain Shams Engineering Journal*, 14(11), 102534.
40. Zhao, Y. L., Tang, J., Huang, H. P., Wang, Z., Chen, T. L., Chiang, C. W., & Chiang, P. C. (2020). Development of iot technologies for air pollution prevention and improvement. *Aerosol and Air Quality Research*, 20(12), 2874-2888.
41. Abimannan, S., El-Alfy, E. S. M., Hussain, S., Chang, Y. S., Shukla, S., Satheesh, D., & Breslin, J. G. (2023). Towards federated learning and multi-access edge computing for air quality monitoring: literature review and assessment. *Sustainability*, 15(18), 13951.

42. Saini, J., Dutta, M., & Marques, G. (2020). Indoor air quality prediction using optimizers: A comparative study. *Journal of Intelligent & Fuzzy Systems*, 39(5), 7053-7069.
43. Khan, A. N., Rizwan, A., Ahmad, R., & Kim, D. H. (2023). An OCF-IoTivity enabled smart-home optimal indoor environment control system for energy and comfort optimization. *Internet of Things*, 22, 100712.
44. Chen, H. C., Putra, K. T., Tseng, S. S., Chen, C. L., & Lin, J. C. W. (2020). A spatiotemporal data compression approach with low transmission cost and high data fidelity for an air quality monitoring system. *Future Generation Computer Systems*, 108, 488-500.
45. Anastasi, G., Bartoli, C., Conti, P., Crisostomi, E., Franco, A., Saponara, S., ... & Vallati, C. (2021). Optimized energy and air quality management of shared smart buildings in the covid-19 scenario. *Energies*, 14(8), 2124.
46. Kausar, A., Ahmad, I., Zhu, T., Shahzad, H., & Eisa, M. H. (2023). Exigency for the control and upgradation of indoor air quality—forefront advancements using nanomaterials. *Pollutants*, 3(1), 123-149.
47. Wiryasaputra, R., Huang, C. Y., Kristiani, E., Liu, P. Y., Yeh, T. K., & Yang, C. T. (2023). Review of an intelligent indoor environment monitoring and management system for COVID-19 risk mitigation. *Frontiers in public health*, 10, 1022055.
48. Nguyen, H. A., & Ha, Q. P. (2022). Wireless sensor network dependable monitoring for urban air quality. *IEEE Access*, 10, 40051-40062.
49. Gololo, M. G. D., Nyathi, C. W., Boateng, L., Nkadameng, E. K., Mckenzie, R. P., Atif, I., ... & Mellado, B. (2024). Review of IoT Systems for Air Quality Measurements Based on LTE/4G and LoRa Communications. *IoT*, 5(4), 711-729.
50. Kamran, M., Salahuddin, H., Kashif, I., Bashir, U., Danish, A., & Rehman, A. U. (2024). Air Quality and Carbon Monoxide Monitoring Using IOT-based System. *Journal of Computing & Biomedical Informatics*, 7(01), 476-483.