Smart Internet of Things

www.siot.reapress.com

Smart. Internet. Things. Vol. 1, No. 3 (2024) 226–243.

Paper Type: Original Article

AI-Driven IoT Solutions for Urban Pollution Monitoring

Chirayil Alex Binu*

School of Computer Science Engineering, KIIT University, Bhubaneswar, India; 2229110@kiit.ac.in.

Citation:

Received: 12 August 2024 *Revised: 09 October 2024* *Accepted: 17 December 2024*	Binu, C. A. (2024). AI-driven IoT solutions for urban pollution monitoring. *Smart internet of things, 1*(3), 226-243.

Abstract

Urban pollution is a growing problem that affects public health, the quality of the environment, and living conditions in cities. Conventional methods of monitoring pollution often do not provide real-time data or predictive insights, which hampers effective responses. The integration of Artificial Intelligence (AI) with the Internet of Things (IoTs) presents an innovative solution for monitoring urban pollution through cutting-edge sensors, data analysis, and forecasting techniques. This paper investigates the structure, application, and success of AI-enhanced IoTs systems for real-time pollution monitoring in urban environments. Results indicate that these systems facilitate proactive management of pollution, enhance urban planning, and increase public engagement, making them vital resources for tackling pollution issues in cities around the globe.

Keywords: Artificial intelligence, Internet of things, Pollution monitoring, Urban environment, Data analysis, Predictive modeling.

1 | Introduction

1.1 | Background

As urban populations grow, so do the challenges associated with maintaining a healthy environment. Urban air pollution has become one of the most pressing public health crises, affecting millions worldwide. Pollutants such as Carbon Monoxide (CO), Sulfur Dioxide (SO_2), Nitrogen Oxides (NOx), and particulate matter (PM2.5 and PM10), largely stemming from vehicle emissions, industrial processes and energy production, are known to contribute to respiratory and cardiovascular illnesses. Moreover, the increased presence of fine particles in the atmosphere exacerbates the risk of chronic diseases and shortens life expectancy [1], [2].

Studies have shown that traditional pollution monitoring networks struggle to capture the complex spatial variability of pollutants in large urban areas. Consequently, the urgent need for real-time, high-resolution

Corresponding Author: 2229110@kiit.ac.in
https://doi.org/10.22105/siot.v1i3.217

monitoring has driven research towards Internet of Things (IoTs)-based monitoring systems, which utilize networks of low-cost, widely distributed sensors [3]–[5].

1.2 | Limitations of Traditional Monitoring Systems

Conventional pollution monitoring systems, typically deployed as stationary ground-based sensors, offer limited spatial coverage due to high installation and maintenance costs. These fixed sensors cannot capture the rapid changes in pollution levels, leading to an underrepresentation of pollution hotspots and delayed responses to pollution events.

These limitations hinder urban authorities' ability to identify and respond to pollution events in real-time, leaving the public exposed to potentially dangerous air quality levels. While useful for broad geographic coverage, satellite-based monitoring lacks the resolution and consistency needed for local monitoring due to its reliance on weather conditions and limited revisit rates [6].

1.3 | The Role of Internet of Things and Artificial Intelligence in Pollution Monitoring

The advent of IoTs and Artificial Intelligence (AI) technologies has transformed pollution monitoring [7]–[9]. When deployed in a high-density network, IoTs sensors capture pollution data at finer spatial and temporal resolutions than traditional systems. AI, particularly Machine Learning (ML) and Deep Learning (DL), enables this data to be analyzed in real-time, uncovering patterns that can predict pollution trends and trigger alerts before pollution reaches hazardous levels [10].

IoTs and AI form a robust framework that enables urban authorities to make informed, data-driven decisions, implement immediate interventions, and ultimately protect public health more effectively.

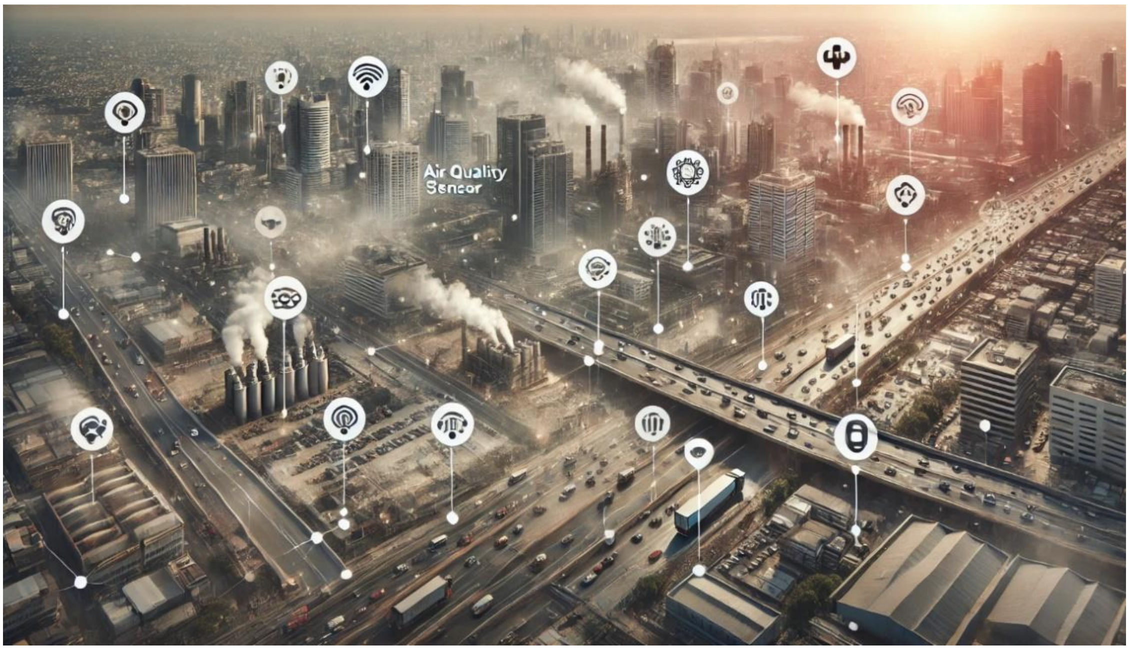

Fig. 1. Diagram of a city with distributed internet of things sensors for pollution monitoring, transmitting data to a central artificial intelligence-driven analysis hub.

This paper investigates how AI-driven IoTs systems are deployed for urban pollution monitoring, emphasizing the system's architecture, the role of AI in data processing, and IoT's potential for expansive and responsive monitoring networks. By integrating real-time pollution monitoring with predictive capabilities, AI-driven IoTs systems can be powerful tools for city planners, policymakers, and citizens to make informed decisions. This paper also discusses the technical and ethical challenges associated with deploying these systems on a large scale and the future opportunities for further innovation in this field.

2 | Literature Review

2.1 | Traditional Pollution Monitoring Techniques

Historically, air quality monitoring relied on ground-based monitoring stations and remote sensing technologies. These traditional methods provide baseline data but fail to capture the intricate variations of air quality in densely populated urban areas. Key limitations include:

I. Limited coverage: traditional monitoring stations are expensive and offer limited spatial coverage. As a result, only a fraction of urban areas is monitored.

II. Delayed reporting: these stations report data at intervals (e.g., hourly or daily), limiting the potential for immediate response to pollution peaks.

Satellite imagery complements ground-based monitoring by providing broader coverage. Still, its lower spatial resolution and sensitivity to weather conditions make it unreliable for tracking pollution in real-time within city environments [11].

2.2 | Advances in Internet of Things for Environmental Monitoring

The IoTs has revolutionized environmental monitoring through affordable and versatile sensors that can be deployed in large numbers across urban areas [12]. This approach overcomes many limitations of traditional systems by:

I. Increasing data granularity: IoTs sensors can be deployed every few hundred meters, capturing highly localized data.

II. Facilitating real-time data collection: IoTs sensors continuously collect and transmit data, enabling near-instantaneous monitoring of air quality changes.

Research shows IoTs sensors effectively track CO_2, PM2.5, and ozone pollutants. However, the data collected by IoTs sensors often contains noise due to environmental interference and sensor drift, which AI algorithms can mitigate [13].

2.3 | Role of Artificial Intelligence in Enhancing Pollution Data Accuracy

AI techniques, including ML and DL [14], are crucial for processing and interpreting the large datasets generated by IoTs sensors. AI models can identify patterns in pollution levels, forecast trends, and perform anomaly detection to flag unusual pollution events. For instance, predictive models may incorporate weather conditions, traffic patterns, and historical pollution data to anticipate pollution spikes, enabling preemptive action.

Additionally, advanced AI algorithms, such as Convolutional Neural Networks (CNNs), can be used for image analysis in satellite data, providing complementary data on pollution hotspots. AI techniques, including ML and DL, are crucial for processing and interpreting large datasets generated by IoTs sensors. AI models can identify patterns in pollution levels, forecast trends, and perform anomaly detection to flag unusual pollution events.

For instance, predictive models may incorporate weather conditions, traffic patterns, and historical pollution data to anticipate pollution spikes, enabling preemptive action. Additionally, advanced AI algorithms, such as CNNs, can be used for image analysis in satellite data, providing complementary data on pollution hotspots.

2.4 | Combining Internet of Things and Artificial Intelligence for Smart Pollution Monitoring

Integrating IoTs with AI enables a new level of predictive analytics in pollution monitoring. By collecting and analyzing data continuously, cities can create pollution models that forecast air quality hours or even days in advance. This allows for preventive actions, such as regulating traffic flow or adjusting industrial activity [15].

Table 1. A comparative table of traditional vs. Internet of things-based monitoring systems, outlining coverage, cost, and data granularity differences.

Feature	Traditional Monitoring Systems	IoTs-based Monitoring Systems
Coverage	Limited range; often confined to specific areas.	Wide coverage; can monitor multiple locations simultaneously.
Cost	High upfront costs for installation and maintenance of physical infrastructure.	Lower upfront costs but may incur ongoing expenses for cloud services and maintenance.
Data granularity	Generally provides aggregated data, often at intervals (e.g., daily, weekly).	Offers high-resolution data with real-time monitoring and the ability to analyze minute-by-minute variations.
Setup complexity	It requires manual setup and maintenance of physical devices; it is often labor-intensive.	It can be quickly deployed with minimal manual intervention; devices are often self-configuring.
Data analysis	Limited analytical capabilities often require manual analysis and reporting.	Advanced analytics using AI and ML, enabling predictive maintenance and trend analysis.
Scalability	Difficult to scale due to physical constraints; often requires significant investment.	Easily scalable by adding more devices and sensors as needed without extensive infrastructure changes.
User accessibility	Data is often less accessible, typically available only to specific personnel.	Data can be accessed remotely via web and mobile applications, enhancing user accessibility.
Maintenance	Regular physical maintenance is required, often leading to downtime.	Remote diagnostics and updates reduce maintenance requirements and downtime.

3 | The Role of Internet of Things in Pollution Monitoring

3.1 | Internet of Things Sensor Networks

The foundation of an IoTs-based pollution monitoring system lies in the sensor network, where devices are strategically positioned to cover various urban environments. Typical sensors include particulate matter detectors, gas, and acoustic sensors, each tailored to capture specific pollutants or environmental parameters. For example, PM2.5 and PM10 sensors monitor particulate matter in the air, while NO_2 and CO sensors measure gaseous pollutants from vehicle exhaust. Therefore, a robust IoTs sensor network provides multi-dimensional pollution data, enabling authorities to understand how pollutants interact and affect urban health [16].

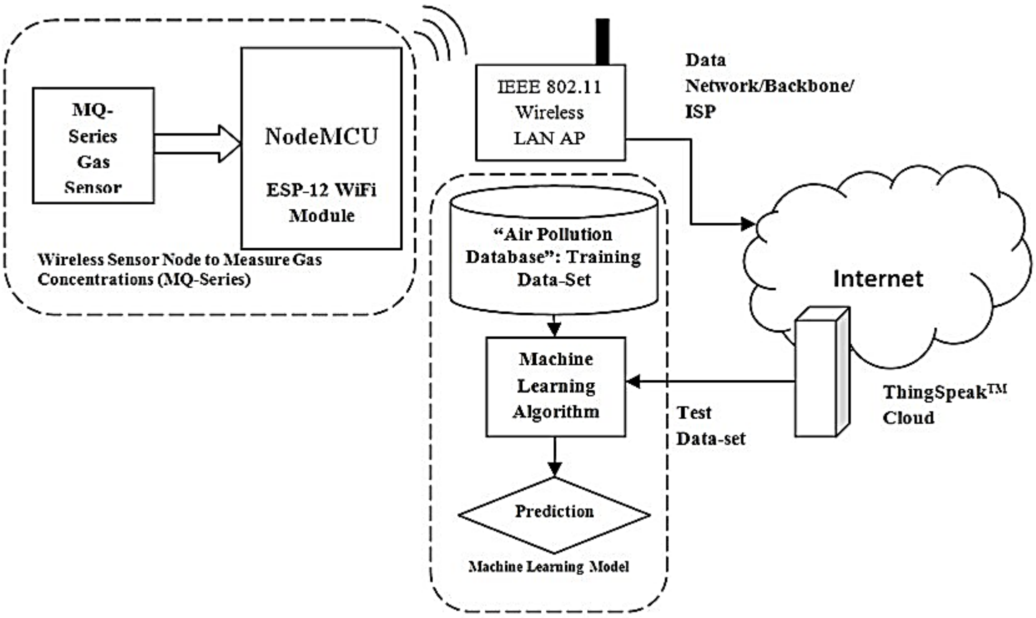

Fig. 2. Internet of things sensor network for urban pollution monitoring.

3.2 | Data Collection and Transmission

Data from IoTs sensors is transmitted in real-time over various wireless communication protocols, such as LoRa, NB-IoTs, and LTE, depending on the network's reach and the device's power constraints. The high data transmission frequency allows for near-instantaneous monitoring of pollution levels, which is essential for timely response in severe pollution episodes. The data is then aggregated in a central database, ready for processing by AI algorithms.

3.3 | Challenges with Internet of Things in Urban Environments

Urban settings pose unique challenges to IoTs networks, such as interference from buildings and other structures, regular maintenance, and sensor calibration to ensure accurate readings. Additionally, managing the large data volumes produced by these sensors presents a data handling and storage challenge, particularly when scaling systems across entire cities [17]–[19].

4 | The Role of Artificial Intelligence in Data Processing and Analysis

4.1 | Data Processing with Artificial Intelligence

AI's role in processing IoTs data involves several stages, including data cleaning, normalization, and integration, which ensure data consistency and reliability. This step is critical, as inaccurate data can lead to misinterpretation and flawed predictions. Additionally, AI efficiently handles vast amounts of data, making it possible to analyze real-time information from thousands of sensors across a city.

4.2 | Predictive Models and Real-time Decision Making

AI models like neural networks and Support Vector Machines (SVMs) can predict pollution levels by analyzing historical data alongside current sensor readings. These predictions inform decision-makers, allowing them to manage potential pollution risks preemptively. Additionally, real-time data feeds enable AI algorithms to generate alerts if pollution levels exceed safe thresholds, improving response times.

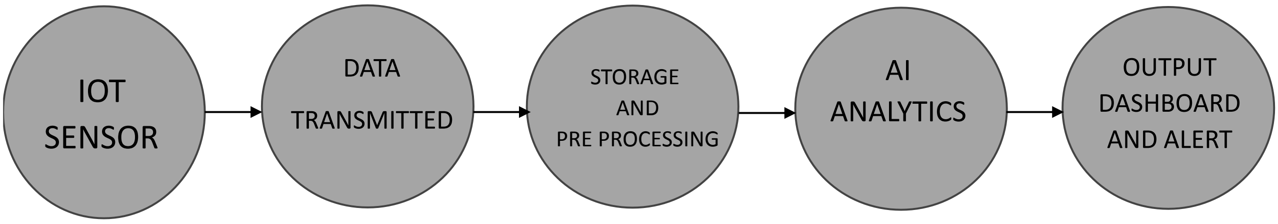

Fig. 3. Data flow in an artificial intelligence-driven internet of things system.

4.3 | Anomaly Detection and Alerts

Anomaly detection algorithms within AI systems can identify unusual pollution patterns and alert relevant authorities. Such systems are essential in urban areas where pollution sources vary, and anomalies can indicate dangerous levels of pollutants that may require immediate action.

5 | System Architecture of Artificial Intelligence-Driven Internet of Things for Pollution Monitoring

The proposed architecture for AI-driven IoTs in urban pollution monitoring consists of several interconnected components designed to work seamlessly to provide real-time, actionable insights. This modular and scalable system architecture allows for easy integration of new sensors, processing nodes, and AI algorithms as the network grows. Below is a detailed breakdown of each subsystem within the architecture:

5.1 | Internet of Things Network and Sensor Deployment

The foundation of the system architecture is a network of IoTs sensors strategically deployed across the urban area. These sensors measure various pollutants and are installed on:

I. Streetlights and traffic signals, where they can capture emissions from vehicles.

II. Public buildings and parks provide a broader assessment of air quality away from roadways.

III. Residential and commercial rooftops help to detect pollution patterns at different altitudes.

Each sensor is designed to monitor specific pollutants, such as carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ozone (O_3), and particulate matter (PM2.5 and PM10). Sensors for environmental factors like temperature, humidity, and wind speed are also included, as these can affect pollutant dispersion.

The sensors communicate wirelessly using Low-Power Wide-Area Network (LPWAN) protocols, like LoRaWAN or NB-IoTs, which provide long-range, energy-efficient connectivity crucial for large-scale urban deployment.

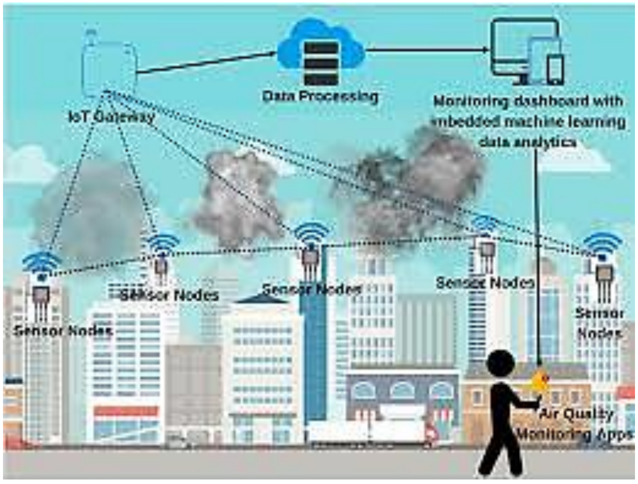

Fig. 4. Diagram of internet of things sensor nodes spread across an urban area with data routed to a central processing system.

5.2 | Edge Processing and Data Preprocessing

Each IoTs sensor node is equipped with basic processing capabilities that allow for edge processing:

I. Initial data filtering: data collected by sensors is filtered at the edge to remove outliers or extreme values, minimizing the need to transfer large amounts of raw data.

II. Data compression: only essential features are transmitted to the cloud, which reduces network bandwidth requirements and storage costs.

Edge processing significantly reduces latency and helps address network congestion by limiting the volume of data transmitted to the central hub. This decentralized processing layer ensures data continuity even in areas with intermittent connectivity.

5.3 | Data Transmission and Communication Protocols

Data from the IoTs nodes is transmitted via LPWAN to a series of regional gateways, which collect, consolidate, and securely transmit the data to the central cloud or server for further processing. Encryption protocols like AES-256 are applied during transmission to maintain data integrity and minimize security risks. The system supports both real-time and batch processing:

I. Real-time data streams: pollutant levels are continuously monitored, with real-time data streams enabling rapid response to pollution events.

II. Batch processing: historical data is aggregated in periodic batches, enabling more complex analytical processes and longitudinal studies without impacting real-time data flow.

5.4 | Centralized Data Processing and Artificial Intelligence Model Integration

The centralized data processing layer aggregates, cleans, and analyzes data from various sensors. This layer consists of cloud servers or local data centers equipped with high-capacity processing units:

I. Data cleansing and normalization: raw data undergoes cleansing to remove duplicates, outliers, or erroneous entries. Normalization is applied to standardize data values, ensuring compatibility across different sensors and locations.

II. Feature extraction and selection: key attributes such as peak pollutant levels, diurnal trends, and meteorological variables are extracted to optimize model accuracy.

The AI models used in this system include:

I. Classification algorithms (e.g., SVMs): used to categorize zones into various air quality levels (e.g., low, medium, high).

II. Regression models (e.g., Random Forest and Linear Regression) predict pollution trends, offering insights on likely future pollution levels in specific locations.

III. DL models (e.g., Long Short-Term Memory (LSTM) networks) analyze temporal patterns in pollution, allowing for forecasting pollution peaks based on past trends.

5.5 | Real-time Analytics Dashboard and Alert System

The processed data is displayed on an analytics dashboard that can be accessed by city officials, researchers, and the public. The dashboard includes:

I. Heat maps of real-time pollution levels across the city, with color-coded zones indicating air quality status.

II. 24-hour predictive models that allow authorities to anticipate pollution spikes and prepare appropriate responses.

III. Alerts and notifications: the system sends automatic alerts if pollution levels exceed set thresholds, allowing rapid action from emergency services, traffic management, and public health departments.

The dashboard is optimized for ease of use, offering options for accessing historical data, setting custom alerts, and generating reports on demand.

Fig. 5. Example dashboard with heatmaps, real-time alerts, and a 24-hour pollution forecast model[1].

5.6 | Data Storage and Long-term Analysis

The architecture includes a data storage system that archives all collected data over extended periods to facilitate historical data analysis. This archive enables long-term studies on pollution trends, seasonal changes' impact, and policy interventions' effectiveness. Advanced analytics, such as predictive modeling and anomaly detection, are periodically applied to this data to identify patterns that inform future urban planning and public health policies.

[1]https://www.inetsoft.com/info/air-quality-dashboard-kpis-and-analytics/

6 | Proposed Methodology

The proposed methodology for implementing an AI-driven IoTs framework for urban pollution monitoring is structured around a multi-stage approach that ensures high-quality data collection, robust analysis, and actionable insights. This methodology is designed to operate in real-time, providing continuous updates and forecasts that enable city authorities to respond promptly to pollution events.

6.1 | Data Collection from Distributed Internet of Things Sensors

The methodology's initial phase focuses on data collection through a dense network of IoTs sensors deployed across key urban locations. Each sensor node collects data on various pollutants, including CO_2, PM2.5, PM10, NOx, and SO_2, and meteorological variables such as temperature, humidity, and wind speed, influencing pollution dispersion patterns. Key steps in data collection:

I. Sensor placement and calibration: sensors are strategically placed near pollution sources (e.g., major roads and industrial zones) and sensitive areas (e.g., schools and hospitals) to maximize data relevance. Regular calibration ensures data accuracy.

II. Real-time data transmission: each sensor has wireless connectivity (e.g., LoRaWAN, NB-IoTs) that enables real-time data transmission to local gateways or directly to the cloud for further processing.

III. Continuous data stream: the data is collected at short intervals (e.g., every 5-10 seconds), providing a near-continuous stream that reflects real-time air quality changes.

Challenges: due to environmental factors, data from IoTs sensors can be noisy, requiring advanced preprocessing techniques to ensure accuracy.

6.2 | Preprocessing and Initial Data Filtering

Raw data collected from IoTs sensors is preprocessed to maintain high data quality before entering the main analytical pipeline. Preprocessing involves several steps to remove errors, reduce noise, and standardize the data. Key preprocessing steps:

I. Data cleansing: erroneous data points caused by sensor drift, environmental interference, or technical faults are filtered out. Statistical techniques like median filtering are used to detect and correct outliers.

II. Data normalization and standardization: pollution readings are normalized to a common scale to ensure compatibility across various sensor types and locations. This step prevents disparities in data ranges from influencing AI model outcomes.

III. Noise reduction: techniques like Kalman filtering and moving averages reduce signal noise and smooth out short-term fluctuations in data.

Tools and techniques: edge computing resources on the sensor nodes allow initial data filtering, reducing network load by transmitting only essential data to the central server.

6.3 | Feature Extraction and Selection

Specific features are extracted from the cleaned data to effectively predict and classify pollution. The feature selection process identifies parameters contributing significantly to pollution events and patterns. Important features include:

I. Pollutant concentrations: pollutants like PM2.5, NO_2, and SO_2 are extracted and tracked to assess pollution trends.

II. Meteorological variables: temperature, wind speed, and humidity levels are included as they significantly affect the dispersion and concentration of airborne pollutants.

III. Temporal features: temporal aspects, such as time and day of the week, are included to capture routine traffic or industrial patterns that may influence air quality.

ML techniques like Principal Component Analysis (PCA) identify and retain only the most relevant features, improving model accuracy and reducing computational costs.

6.4 | Artificial Intelligence Model Training and Deployment

The proposed methodology's core is the AI model, which analyzes preprocessed data and predicts pollution levels. Multiple machines and DL models are trained on historical pollution and meteorological data to ensure robust predictions. Model training process:

Data splitting

The preprocessed dataset is split into training and testing subsets to ensure that models generalize well on new data.

Training algorithms

I. Regression models: algorithms like Random Forest and Linear Regression are used to predict pollutant levels continuously.

II. Classification models: SVMs and Logistic Regression classify zones by pollution level (e.g., low, moderate, high).

III. Time-series forecasting: for predictive capabilities, LSTM networks analyze temporal patterns and provide short-term forecasts of pollution trends.

Cross-validation and hyperparameter tuning

Cross-validation is used alongside hyperparameter tuning to find the best configurations for each model and enhance model accuracy. Techniques like Grid Search and Random Search are applied to optimize parameters such as learning rate, depth of trees, and regularization terms.

Deployment strategy: the trained models are deployed in the cloud, continuously analyzing incoming data. Edge computing may also be used for lighter models to perform initial analysis closer to the data source, reducing latency.

6.5 | Prediction and Real-time Analysis

Once the models are deployed, they begin analyzing real-time data streams. This phase involves both predictive analysis and anomaly detection, allowing the system to forecast pollution spikes and notify authorities of unusual patterns. Steps in real-time analysis:

I. Short-term pollution forecasting: AI models predict pollutant levels over the next few hours, offering timely insights into impending pollution peaks.

II. Anomaly detection: ML algorithms identify unusual pollution events or spikes that deviate from regular patterns, potentially signaling events like industrial emissions or traffic congestion.

The analysis provides a dynamic view of pollution levels, allowing urban authorities to monitor air quality across different zones in near real-time.

6.6 | Alert Mechanism and Decision Support

An alert system notifies relevant stakeholders of potential pollution events based on predictions and anomaly detections. This alert mechanism includes the following:

I. Threshold-based alerts: predefined thresholds for each pollutant trigger alerts when exceeded. For instance, if PM2.5 levels surpass safe limits, environmental authorities automatically send an alert.

II. Dynamic risk assessment: alerts are prioritized based on factors such as the pollutant type, concentration, and exposure risk to sensitive locations (e.g., schools and hospitals).

III. Public notifications: in severe cases, notifications can be sent to the public via mobile apps or digital signage in public spaces, advising people to limit outdoor activities or use protective gear.

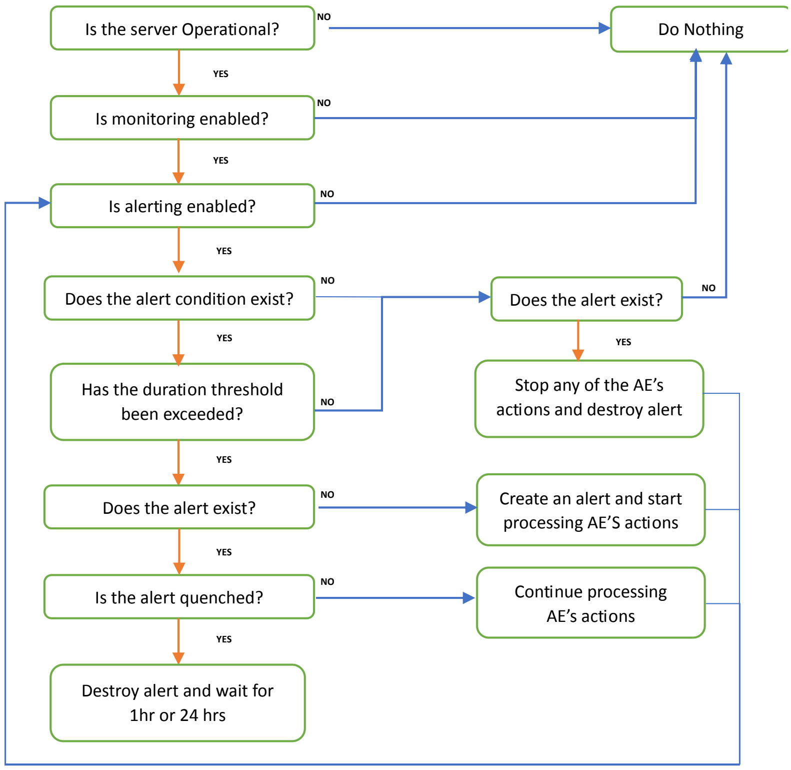

Fig. 6. Flowchart detailing end-to-end methodology, from data collection to alert generation.

6.7 | Visualization and User Dashboard

The methodology culminates in a user-friendly dashboard that visualizes the system's outputs, making it accessible to city planners, environmental agencies, and the public. This visualization provides a clear overview of pollution trends and forecasts, enabling informed decision-making. Dashboard features:

I. Heatmaps: real-time heatmaps display pollution levels across different zones in the city, with color codes representing air quality status (e.g., good, moderate, unhealthy).

II. Historical data trends: users can view historical trends to identify areas with recurring pollution issues and understand long-term patterns.

III. Forecast models: predictive models show expected pollution levels for the upcoming hours, enabling proactive responses to anticipated pollution spikes.

6.8 | Long-term Data Analysis and Model Improvement

The system stores collected and analyzed data for long-term studies, enabling city authorities to track the impact of pollution management strategies and assess seasonal trends. This archived data allows for periodic model retraining to ensure accuracy over time. Steps for model improvement:

I. Model retraining: periodic retraining of AI models using newly collected data keeps predictions accurate as urban dynamics change.

II. Integration of new features: as new data sources and sensor types become available, additional features can be integrated into the models to improve predictive capabilities.

III. Comparative analysis: regular comparisons between predicted and actual pollution data help refine model parameters and highlight areas for improvement.

7 | Challenges and Limitations

7.1 | Data Quality and Reliability

A key challenge in deploying IoTs-based pollution monitoring systems is ensuring the quality and reliability of data collected from the sensors. IoTs sensors, though cost-effective and scalable, can be sensitive to environmental conditions, leading to issues such as:

I. Sensor drift: IoTs sensors may show reduced accuracy over time due to wear or environmental exposure, which can lead to data drift. For example, temperature fluctuations or high humidity can affect the performance of gas sensors, resulting in inconsistent readings.

II. Noise in data: environmental noise, electrical interference, and signal degradation can introduce erroneous data points. Filtering out these anomalies is essential to maintaining accurate real-time monitoring and prediction, but this process requires advanced algorithms, which can be computationally expensive.

To address these challenges, regular recalibration and automated data-cleaning algorithms are necessary. AI-driven filtering methods, such as Kalman filtering or robust regression, can help, but these methods must be carefully tuned to avoid filtering out genuine, unusual pollution spikes.

7.2 | Scalability and Infrastructure Requirements

Scaling an IoTs network to cover a large urban area introduces substantial infrastructure and resource demands:

I. Network infrastructure: A dense deployment of IoTs sensors across a city requires robust network infrastructure to manage data transmission without excessive latency or data loss. LPWAN and 5G can support data transfer from many sensors, but implementing these technologies is costly and requires substantial planning.

II. Data storage and processing: the continuous data streams from numerous IoTs nodes generate vast data. Managing, storing, and processing this data in real time places considerable demands on cloud infrastructure, especially in larger cities. Advanced data compression techniques and edge computing solutions can mitigate challenges by processing data closer to the source, reducing the strain on central servers.

III. Energy efficiency and maintenance: IoTs sensors deployed across a city must function on minimal power, often relying on battery sources. As urban IoTs networks scale up, ensuring consistent power supply and minimizing maintenance needs becomes crucial. Smart cities with integrated power sources or solar panels can alleviate these constraints, but implementing them increases initial setup costs.

7.3 | Privacy and Security Concerns

Since IoTs devices transmit data wirelessly over public networks, maintaining data privacy and security is a critical concern, especially when monitoring in urban areas where data can reveal sensitive location and environmental information.

I. Data encryption: transmitted data needs to be encrypted to prevent unauthorized access. Encryption protocols for IoTs networks must balance security with low power and processing requirements. Lightweight encryption methods, like Elliptic Curve Cryptography (ECC), are commonly used but may still introduce data transmission latency.

II. Data anonymization: location data is often essential for monitoring in urban settings; however, anonymizing this data to avoid infringing on privacy rights is challenging. Techniques such as differential privacy can protect individual privacy by adding noise to the data, but they must be carefully managed to prevent data distortion.

III. Cybersecurity threats: IoTs devices are vulnerable to various cyberattacks, such as data spoofing, where false data is injected into the network, or Distributed Denial of Service (DDoS) attacks that can disrupt entire networks. Implementing cybersecurity measures in IoTs systems is essential but often overlooked due to cost or technical constraints.

7.4 | Environmental and Urban Challenges

Urban environments introduce specific challenges that can impact the accuracy and reliability of IoTs sensors:

I. Variability in pollution sources: urban pollution is dynamic, with varying sources such as vehicle emissions, industrial discharges, and construction activities. These sources can create localized pollution spikes that are difficult for a static monitoring network to capture consistently. An AI-driven approach is required to adapt to these fluctuations and adjust sensor sensitivity based on identified hotspots.

II. Interference from structures: buildings, walls, and other physical structures can interfere with signal transmission between IoTs devices, reducing data quality. Care sensor placement planning is required to ensure continuous data flow, especially in dense urban areas.

III. Traffic and human interference: pedestrian and vehicle traffic may obstruct or damage sensors. Furthermore, heavy vehicular movement near sensors can generate additional particles or gases, complicating readings.

The system must adapt to address these environmental challenges in real-time, using dynamic sensor positioning and adaptive AI models that adjust based on environmental inputs and sensor feedback. However, such adaptability increases system complexity and demands higher computational resources.

7.5 | Algorithm Complexity and Computational Requirements

AI models used in pollution monitoring require substantial computational resources. Complex AI algorithms, such as DL models, often involve large datasets, necessitating powerful processors, high memory, and sometimes even specialized hardware like GPUs. This introduces several specific challenges:

I. Processing and latency: high computational loads can result in data processing delays, which limit the system's ability to provide real-time alerts. Optimizing the processing pipeline with distributed computing or using more efficient algorithms (e.g., lightweight neural networks) is necessary, but it still requires careful balancing between accuracy and speed.

II. Model training and updating: AI models must be regularly updated with new data to maintain accuracy in predicting pollution levels. However, frequent retraining can strain resources, particularly if models are large and computationally intensive. Edge computing, where certain processing tasks are performed at the device level, can alleviate central processing load but requires advanced hardware in the IoTs devices.

To manage these complexities, hybrid models that combine lightweight processing at the edge with more advanced analytics at a centralized location can be employed. However, this configuration adds to the system's setup and operational costs.

7.6 | Societal and Regulatory Challenges

Urban pollution monitoring often faces societal and regulatory obstacles, as monitoring and controlling pollution involve coordination across various government agencies, private sector partners, and the public. Key challenges include:

I. Data sharing and ownership: sharing pollution data across organizations can be complex, particularly in cities with privatized data networks. Questions around data ownership, access rights, and liability in cases of inaccurate data can create legal and operational barriers.

II. Regulatory standards for IoTs sensors: ensuring that IoTs sensors comply with national and international data accuracy and reliability standards is challenging, especially as standards evolve. Cities must adopt regulatory frameworks that mandate calibration and maintenance schedules, which add to operational costs.

III. Public acceptance and awareness: public acceptance is crucial, as IoTs sensors may be viewed as invasive. Building public trust involves transparent data-sharing practices, privacy safeguards, and regular communication on the system's benefits.

A proactive approach to regulatory engagement, including setting up clear data ownership policies and public awareness campaigns, is essential for long-term success. However, such initiatives require coordination and financial support from local governments and private sector partners.

8 | Future Work

The field of AI-driven IoTs solutions for urban pollution monitoring is still evolving, and numerous potential advancements can further enhance such systems' accuracy, scalability, and usability. Here, we discuss several promising areas for future development:

8.1 | Advanced Sensor Technology

Multi-functional Sensors

IoTs sensors typically focus on single pollutants (e.g., CO or PM2.5). Future developments may include multi-functional sensors capable of measuring multiple pollutants simultaneously. This would reduce deployment costs and improve the spatial resolution of data.

Self-calibrating sensors

One of the main challenges with IoTs sensors is calibration drift over time. New technologies, such as self-calibrating sensors, could improve data quality by automatically adjusting to environmental changes, minimizing the need for manual recalibration.

8.2 | Enhanced Artificial Intelligence Models and Predictive Capabilities

Deep learning for complex pattern recognition

AI models can be expanded to use DL architectures like CNNs for image-based pollution detection or Recurrent Neural Networks (RNNs) for time-series forecasting. Such models could improve predictive capabilities, particularly for dynamic or highly localized pollution patterns.

Federated learning for decentralized data processing

Federated learning enables AI models to learn from data stored on individual IoTs devices without transferring data to a central server. This approach could reduce network congestion, enhance privacy, and enable more efficient, large-scale deployments. Federated learning could allow pollution models to adapt locally while sharing generalized insights with the central system.

8.3 | Integration with Other Urban Systems

Traffic management systems

Integrating pollution monitoring systems with urban traffic management can lead to more comprehensive control over pollution sources. For instance, traffic could be rerouted in areas with high NO_2 levels to reduce emissions and mitigate pollution.

Smart city ecosystem expansion

A broader integration of pollution data into smart city ecosystems, such as public health, emergency response, and transportation systems, would enable coordinated responses to pollution events. Additionally, the data could be shared with public transit systems to adjust routes or schedules based on air quality levels.

8.4 | Real-time Public Engagement and Health Alerts

Mobile applications for public awareness

In addition to dashboards for city authorities, mobile applications could be developed so that the public can access real-time air quality information, receive alerts, and receive personalized health advice based on pollution levels in their vicinity.

Personalized health risk assessment

Future systems could offer tailored health risk assessments, especially for vulnerable populations such as children, the elderly, and those with respiratory conditions. By using health data integrated with pollution data, these alerts could help residents make informed decisions about outdoor activities during high-pollution periods.

8.5 | Improved Data Security and Privacy Protocols

As IoTs deployments scale up, ensuring data security and privacy is paramount. Future developments could involve:

I. Blockchain for data integrity: blockchain technology offers a secure way to manage and verify data from IoTs sensors, reducing the risk of data tampering and ensuring transparency across the data lifecycle.

II. Privacy-enhanced algorithms: as data privacy regulations evolve, future systems should implement privacy-preserving algorithms such as homomorphic encryption and differential privacy. These techniques allow for data analysis without exposing sensitive information, ensuring that personal data remains secure.

8.6 | Scaling and Cross-border Collaborations

Regional and global network integration

Cities worldwide face similar pollution challenges yet often operate in isolation. Developing a standardized framework for pollution monitoring would allow for cross-city comparisons, data sharing, and collaborative efforts. This could also facilitate large-scale studies of pollution patterns, paving the way for international policies on air quality.

Cross-sector partnerships

Collaborations with industries, academic institutions, and environmental organizations could drive innovation in sensor technologies, AI models, and regulatory practices. Such partnerships would help overcome the technical and financial challenges of deploying and maintaining large-scale pollution monitoring systems.

8.7 | Sustainability and Low-energy Solutions

Solar-powered sensors

As urban IoTs networks expand, sustainable power sources are essential for long-term functionality. Solar-powered sensors and energy-efficient designs could extend the lifespan of IoTs devices, reducing the network's environmental impact and operational costs.

Energy-efficient artificial intelligence algorithms

Developing lightweight AI models that consume less power can ensure the system's long-term sustainability. Techniques like model pruning, quantization, and using energy-efficient hardware such as neuromorphic processors are promising areas for future research.

9 | Conclusion

This paper presents a robust framework that integrates the IoTs and AI to address the challenges of urban pollution monitoring. Through a detailed examination of existing methods and limitations, we proposed an

AI-driven IoTs solution capable of continuously collecting, analyzing, and predicting air quality data at a granular level. This system significantly improves over traditional pollution monitoring methods, providing city officials with timely and actionable insights that facilitate more effective pollution control measures.

9.1 | Summary of Contributions

Enhanced data collection

The framework deploys a network of IoTs sensors across diverse urban locations to provide real-time, high-resolution data on multiple pollutants. This allows for more precise detection of pollution hotspots and a better understanding of pollution patterns.

Predictive analytics with artificial intelligence

Integrating AI models like ML and DL enables the system to process vast amounts of data and produce short-term pollution forecasts. These predictive capabilities are vital for urban management, allowing authorities to proactively mitigate pollution spikes and respond more effectively to environmental emergencies.

Real-time alerts and user accessibility

The system ensures that city authorities and the public are informed of critical pollution levels through automated alerts and a user-friendly dashboard. This real-time access to data empowers decision-makers to take immediate action, reducing health risks for vulnerable populations and minimizing the adverse effects of pollution on urban communities.

9.2 | Reflections on Challenges and Solutions

This research highlights the inherent challenges of implementing a large-scale IoTs and AI-based monitoring system in urban areas, such as data accuracy, network scalability, and security concerns. The framework addresses these challenges by incorporating edge processing, encryption protocols, and advanced data-cleaning techniques. It demonstrates the feasibility of a scalable solution that maintains data quality and privacy.

Implications for smart city development and public health

The proposed AI-IoTs solution aligns with broader smart city objectives, leveraging technology to create healthier, more sustainable urban environments. The system's capacity to deliver real-time insights into air quality provides urban planners and policymakers with a data-driven foundation to:

I. Implement effective pollution control measures: short-term predictions enable temporary measures like traffic regulation, while long-term data informs strategies like green space development and industrial regulation.

II. Raise public awareness: with real-time public alerts and data visualizations, citizens gain a clearer understanding of pollution risks, which can encourage responsible environmental practices and foster community-driven sustainability initiatives.

III. Support public health initiatives: access to accurate pollution data helps health officials advise at-risk populations and implement targeted health interventions, reducing hospitalizations related to air quality.

9.3 | Future Research and Development

While the current framework demonstrates significant potential, continued research is essential for addressing its limitations and expanding its functionality. Areas for future work include integrating federated learning to reduce data transmission load, deploying self-calibrating multi-functional sensors, and creating mobile applications for direct public engagement. These advancements will strengthen the system's utility and adaptability in diverse urban settings.

9.4 | Concluding Remarks

In conclusion, combining IoTs and AI in this framework provides a transformative approach to urban pollution monitoring. As cities face increasing urbanization and climate change pressures, such technology-driven solutions are essential to safeguarding public health and promoting sustainable urban growth. The proposed AI-IoTs solution represents an important step toward a future where cities can dynamically monitor and manage air quality, creating safer, healthier, and more resilient communities for future generations.

Data Availability

The data used and analyzed during the current study are available from the corresponding author upon reasonable request.

References

[1] Meo, S. A., Salih, M. A., Alkhalifah, J. M., Alsomali, A. H., & Almushawah, A. A. (2024). Effect of air pollutants particulate matter PM(2.5), PM(10), carbon monoxide (CO), nitrogen dioxide (NO(2)), sulfur dioxide (SO(2)), and ozone (O(3)) on fractional exhaled nitric oxide (FeNO). *Pakistan journal of medical sciences*, *40*(8), 1719–1723. https://doi.org/10.12669/pjms.40.8.9630

[2] Meo, S. A., Salih, M. A., Al-Hussain, F., Alkhalifah, J. M., Meo, A. S., & Akram, A. (2024). Environmental pollutants PM2.5, PM10, carbon monoxide (CO), nitrogen dioxide (NO2), sulfur dioxide (SO2), and ozone (O3) impair human cognitive functions. *European review for medical and pharmacological sciences*, *28*(2), 789–796. https://doi.org/10.26355/eurrev_202401_35079

[3] Felici-Castell, S., Segura-Garcia, J., Perez-Solano, J. J., Fayos-Jordan, R., Soriano-Asensi, A., & Alcaraz-Calero, J. M. (2023). AI-IoT low-cost pollution-monitoring sensor network to assist citizens with respiratory problems. *Sensors*, *23*(23), 9585. https://doi.org/10.3390/s23239585

[4] Popescu, S. M., Mansoor, S., Wani, O. A., Kumar, S. S., Sharma, V., Sharma, A., Arya, V. M, Kirkham, M. B., Hou, D., Bolan, N., & Chung, Y. S. (2024). Artificial intelligence and IoT driven technologies for environmental pollution monitoring and management. *Frontiers in environmental science*, *12*, 1336088. https://doi.org/10.3389/fenvs.2024.1336088

[5] Mohapatra, H. (2021). *Smart city with wireless sensor network: networking of smart city applications*. Kindle.

[6] Huynh, T. L., Fakprapai, S., & Nguyen, T. K. O. (2020). Air quality monitoring with focus on wireless sensor application and data management. In *TORUS 3-toward an open resource using services: cloud computing for environmental data* (pp. 17–40). Wiley online library. https://doi.org/10.1002/9781119720522.ch2

[7] Shukla, P. K., Singh, A., Shaikh, N., Sharma, A., & Singh, R. (2023). Air pollution monitoring by indulging AI and iot for environmental protection. *2023 3rd international conference on pervasive computing and social networking (ICPCSN)* (pp. 1161–1165). IEEE. https://doi.org/10.1109/ICPCSN58827.2023.00197

[8] Alahi, M. E. E., Sukkuea, A., Tina, F. W., Nag, A., Kurdthongmee, W., Suwannarat, K., & Mukhopadhyay, S. C. (2023). Integration of IoT-enabled technologies and artificial intelligence (AI) for smart city scenario: recent advancements and future trends. *Sensors*, *23*(11), 5206. https://doi.org/10.3390/s23115206

[9] Arabelli, R., Boddepalli, E., Buradkar, M., Goriparti, N. V. S., & Chakravarthi, M. K. (2024). IoT-enabled environmental monitoring system using AI. *2024 international conference on advances in computing, communication and applied informatics (ACCAI)* (pp. 1–6). IEEE. https://doi.org/10.1109/ACCAI61061.2024.10602131

[10] Karnati, H. (2023). IoT-based air quality monitoring system with machine learning for accurate and real-time data analysis. *ArXiv preprint arxiv:2307.00580*. https://doi.org/10.48550/arXiv.2307.00580

[11] Yu, T., Wang, W., Ciren, P., & Sun, R. (2018). An assessment of air-quality monitoring station locations based on satellite observations. *International journal of remote sensing*, *39*, 1–16. http://dx.doi.org/10.1080/01431161.2018.1460505

[12] Katie, B. (2024). Internet of things (IoT) for environmental monitoring. *International journal of computing and engineering*, *6*(3), 29–42. http://dx.doi.org/10.47941/ijce.2139

[13] Mohapatra, H. & Rath, A. K. (2021). Designing of fault-tolerant models for wireless sensor network-assisted smart city applications. In *intelligent technologies: concepts, applications, and future directions* (pp. 25-43). Singapore: springer nature singapore. https://doi.org/10.1007/978-981-99-1482-1_2

[14] Gomathi, G., Emilyn, J. J., Thamburaj, A. S., & D, V. K. (2022). Real time air pollution prediction in urban cities using deep learning algorithms and iot. *2022 7th international conference on communication and electronics systems (ICCES)* (pp. 340–343). IEEE. https://doi.org/10.1109/ICCES54183.2022.9835991

[15] Kataria, A., & Puri, V. (2022). AI-and IoT-based hybrid model for air quality prediction in a smart city with network assistance. *IET networks*, *11*(6), 221–233. https://doi.org/10.1049/ntw2.12053

[16] Bobulski, J., Szymoniak, S., & Pasternak, K. (2024). An IoT system for air pollution monitoring with safe data transmission. *Sensors*, *24*(2), 445. https://doi.org/10.3390/s24020445

[17] Hudson-Smith, A., Wilson, D., Gray, S., & Dawkins, O. (2021). Urban IoT: Advances, challenges, and opportunities for mass data collection, analysis, and visualization. In *urban informatics* (pp. 701–719). Springer, singapore. http://dx.doi.org/10.1007/978-981-15-8983-6_38

[18] Rafiq, I., Mahmood, A., Razzaq, S., Jafri, S. H. M., & Aziz, I. (2023). IoT applications and challenges in smart cities and services. *The journal of engineering*, *2023*(4), e12262. https://doi.org/10.1049/tje2.12262

[19] Mukherjee, M., Adhikary, I., Mondal, S., Mondal, A. K., Pundir, M., & Chowdary, V. (2017). A vision of iot: applications, challenges, and opportunities with dehradun perspective. *Proceeding of international conference on intelligent communication, control and devices* (pp. 553–559). Singapore: springer singapore. https://doi.org/10.1007/978-981-10-1708-7_63