

Drone-Based Pollution Detection Using IoT Sensors and Machine Learning

Dr.Yogeshvari Makwana¹, Namaswini Padhy², Zarna Shah³, Sapna Kataria⁴, Chaitali Bhoi⁵, Shivani Sompura⁶

¹Sankalchand Patel College of Engineering, Sankalchand Patel University, Visnagar, Gujarat, India.

²Sankalchand Patel College of Engineering, Sankalchand Patel University, Visnagar, Gujarat, India.

³College of Technology, Silver Oak University, Ahmedabad, Gujarat, India.

⁴Computer Science and Engineering, Institute of Technology Nirma University, Ahmedabad, Gujarat, India.

⁵Computer Science and Engineering, The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat, India.

⁶Computer Engineering, Indus Institute of Technology & Engineering, Indus University, Ahmedabad, Gujarat, India.

Abstract: Air pollution remains one of the most critical environmental and public health challenges worldwide, driven by rapid urbanization, industrialization, and increasing vehicular emissions. Accurate and high-resolution air quality monitoring is essential for understanding pollution dynamics and implementing effective mitigation strategies. Conventional air quality monitoring systems primarily rely on fixed ground-based stations, which are costly, provide limited spatial coverage, and fail to capture vertical and localized variations in pollutant concentrations. To address these limitations, this study proposes a drone-based pollution detection system that integrates Internet of Things (IoT)-enabled sensors with machine learning-based data analytics for real-time and adaptive air quality monitoring. The proposed system employs an unmanned aerial vehicle (UAV) equipped with lightweight environmental sensors capable of measuring key air pollutants, including particulate matter (PM_{2.5} and PM₁₀), carbon monoxide (CO), nitrogen dioxide (NO₂), along with temperature and relative humidity. The UAV collects geo-referenced pollution data at varying altitudes and locations, enabling three-dimensional air quality assessment. Sensor data are transmitted wirelessly to a cloud-based platform using IoT communication protocols, where preprocessing techniques such as noise filtering, normalization, and outlier removal are applied to enhance data reliability. Machine learning algorithms are utilized to analyze both historical and real-time sensor data for pollution prediction and hotspot identification. Multiple supervised learning models, including linear regression, support vector machines, and random forest regression, are evaluated to assess predictive performance. Experimental results indicate that the random forest model achieves superior accuracy due to its ability to model nonlinear relationships among environmental parameters and handle sensor variability. The trained model enables short-term pollution forecasting and supports real-time alert generation when pollutant concentrations exceed predefined safety thresholds. Field experiments conducted in urban and industrial environments demonstrate that the proposed UAV-based system effectively captures spatial and vertical pollution variations that are often overlooked by conventional monitoring stations. The system provides higher spatial resolution, rapid deployment capability, and improved adaptability compared to static air quality monitoring infrastructure. Visualization of processed data through pollution heat maps and dashboards facilitates intuitive interpretation and decision-making for environmental authorities and policymakers. The proposed drone-based pollution detection framework offers a cost-effective, scalable, and flexible solution for modern air quality monitoring applications. Potential use cases include smart city environmental management, industrial emission surveillance, disaster response during hazardous gas leaks, and environmental impact assessment. The integration of UAV technology, IoT architecture, and machine learning analytics represents a significant step toward intelligent and responsive environmental monitoring systems. Future work will focus on multi-drone coordination, adaptive flight path optimization, and the integration of deep learning models to further enhance prediction accuracy and system autonomy.

Keywords: Air Pollution Monitoring, IoT, Machine Learning, Smart Cities, Environmental Sensing Drone-based pollution detection, Unmanned aerial vehicle (UAV), Particulate matter (PM_{2.5}, PM₁₀)

1. Introduction

Air pollution has become one of the most serious environmental and public health challenges worldwide, driven by rapid urbanization, industrial growth, and increased dependence on fossil fuels. The presence of harmful air pollutants such as particulate matter (PM_{2.5} and PM₁₀), carbon monoxide (CO), nitrogen dioxide (NO₂), and other toxic gases has been strongly linked to respiratory illnesses, cardiovascular diseases, and premature mortality. In addition to its impact on human health, air pollution contributes to climate change, ecosystem degradation, and reduced quality of life, particularly in densely populated urban and industrial regions.



Consequently, accurate and continuous air quality monitoring has become a critical requirement for environmental management and policy formulation.

Traditional air quality monitoring systems predominantly rely on fixed ground-based monitoring stations equipped with high-precision instruments. Although these stations provide reliable long-term measurements, they suffer from several limitations. High installation and maintenance costs restrict their deployment to a limited number of locations, resulting in sparse spatial coverage. Furthermore, fixed stations are unable to capture localized pollution sources or vertical variations in pollutant concentrations, which are increasingly important in complex urban environments with high-rise buildings, traffic congestion, and industrial clusters. As a result, conventional monitoring approaches often fail to represent the true spatial and temporal dynamics of air pollution.

Recent advancements in unmanned aerial vehicles (UAVs), sensor miniaturization, wireless communication, and artificial intelligence have enabled new approaches to environmental monitoring. UAVs, commonly referred to as drones, offer unique advantages such as three-dimensional mobility, rapid deployment, and the ability to access hazardous or hard-to-reach areas. When equipped with lightweight air quality sensors, drones can collect pollution data at multiple altitudes and locations, providing high-resolution spatial information that complements existing ground-based monitoring networks.

The integration of the Internet of Things (IoT) further enhances the capabilities of UAV-based monitoring systems. IoT-enabled sensors facilitate real-time data acquisition, wireless transmission, and cloud-based storage, allowing continuous monitoring and remote accessibility. However, the large volume of heterogeneous data generated by such systems necessitates intelligent data processing techniques to extract meaningful insights. Machine learning algorithms have emerged as powerful tools for analyzing environmental data, enabling pollution prediction, trend analysis, and anomaly detection. By learning complex relationships among environmental parameters, machine learning models can improve the accuracy and reliability of air quality assessment.

In this context, this paper proposes a drone-based pollution detection framework that integrates IoT sensors and machine learning techniques for real-time air quality monitoring and analysis. The proposed system aims to overcome the limitations of traditional monitoring approaches by providing flexible deployment, enhanced spatial and vertical resolution, and predictive intelligence. The main contributions of this work include the design of a UAV-mounted IoT sensor platform, the implementation of a cloud-based data processing architecture, and the application of machine learning models for pollution prediction and hotspot identification.

2. Related Work

Air quality monitoring has been an active area of research for several decades, with traditional approaches relying primarily on fixed ground-based monitoring stations. These stations are equipped with high-precision analyzers and provide accurate long-term measurements of atmospheric pollutants. However, due to their high cost, complex infrastructure requirements, and limited deployment density, several studies have highlighted their inability to capture fine-grained spatial and vertical variations in air pollution, particularly in urban and industrial environments.

To overcome these limitations, researchers have explored the use of wireless sensor networks (WSNs) for distributed air quality monitoring. Low-cost sensor nodes deployed across different locations enable wider spatial coverage and real-time data collection. While WSN-based approaches reduce cost and improve coverage, they remain constrained by fixed installation locations and limited vertical sensing capability. Additionally, sensor calibration, power consumption, and data reliability remain major challenges in large-scale WSN deployments.

With advancements in unmanned aerial vehicle (UAV) technology, drone-based air quality monitoring has gained significant attention in recent years. Several studies have demonstrated the feasibility of mounting lightweight gas and particulate matter sensors on UAV platforms to measure pollutants such as CO₂, PM_{2.5}, PM₁₀, and NO₂. UAV-based systems offer enhanced mobility, allowing measurements at different altitudes and locations, which is particularly useful for identifying pollution plumes and localized emission sources. However, early UAV-based approaches primarily focused on data collection and visualization, with limited emphasis on intelligent data analysis and prediction.

The integration of Internet of Things (IoT) technologies has further enhanced UAV-based monitoring systems. IoT architectures enable seamless sensor integration, real-time data transmission, and cloud-based storage and processing. Several researchers have proposed IoT-enabled UAV frameworks for environmental monitoring, emphasizing real-time accessibility and scalability. Despite these advancements, many existing systems rely on threshold-based analysis or basic statistical methods, which are insufficient for capturing complex pollution dynamics influenced by multiple environmental factors.

Machine learning techniques have recently emerged as powerful tools for air quality analysis and prediction. Supervised learning models such as linear regression, support vector machines, decision trees, and random forest algorithms have been applied to predict pollutant concentrations based on meteorological and historical data. More advanced approaches, including deep learning and neural networks, have also been explored for long-term air quality forecasting. While these models demonstrate promising results, most studies utilize data collected from static ground stations or publicly available datasets, limiting their ability to model localized and vertical pollution variations.

A limited number of studies have attempted to combine UAV-based data collection with machine learning-based analytics. These works demonstrate improved spatial resolution and predictive capability; however, challenges such as real-time processing, sensor noise, system scalability, and practical deployment remain open research issues. Furthermore, comparative evaluation of multiple machine learning models using real-time UAV-collected data is often lacking.

In contrast to existing work, the proposed system integrates UAV technology, IoT-based data acquisition, and machine learning analytics into a unified framework for real-time pollution monitoring and prediction. By combining mobile three-dimensional sensing with intelligent data analysis, the proposed approach aims to provide higher accuracy, adaptability, and practical applicability for modern air quality monitoring scenarios.

3. System Architecture

The proposed drone-based pollution detection framework integrates UAV technology, IoT-enabled sensors, and machine learning analytics into a unified system for real-time air quality monitoring. The system is designed to provide high spatial and vertical resolution, flexible deployment, and predictive insights, overcoming the limitations of conventional ground-based monitoring networks. The architecture is composed of four main modules: the UAV platform, the sensor module, data transmission and cloud processing, and the machine learning analysis layer. Each module is described in detail below.

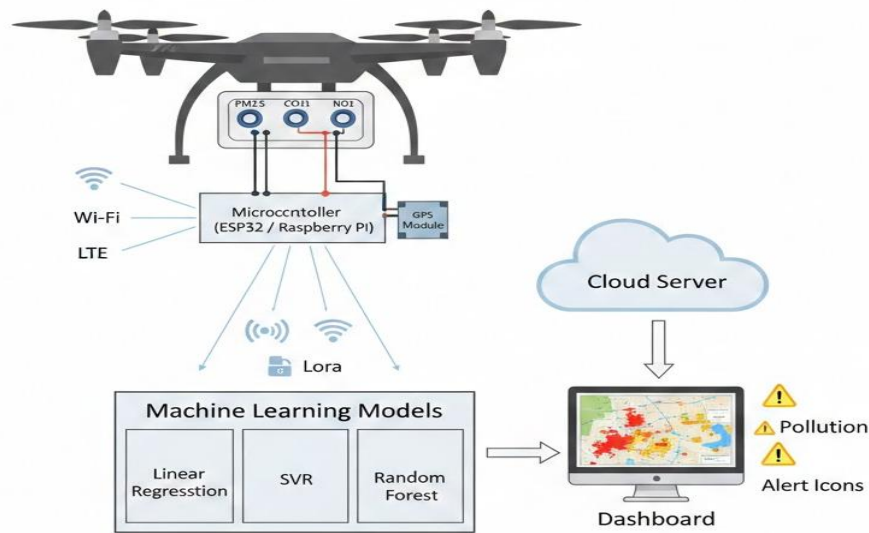


Figure 1: Overall System Architecture of Drone-Based Pollution Detection Using IoT Sensors and Machine Learning.

3.1 UAV Platform

The UAV serves as the primary data collection platform and provides mobility to monitor air quality at various locations and altitudes. A quadcopter drone is chosen for its stability, maneuverability, and payload capacity. The UAV is equipped with:

- **Flight Controller:** Pixhawk or Arduino-based controller for precise navigation and autonomous flight operations.
- **GPS Module:** Provides geo-referenced coordinates to tag sensor data with accurate location information.
- **Power System:** Lithium-polymer (Li-Po) batteries providing sufficient flight endurance (~20–30 minutes per flight).

- Autonomous Flight Software: Capable of following predefined flight paths and altitude profiles while avoiding obstacles.

The UAV can operate in both autonomous and manual modes, enabling flexibility in data collection over urban, industrial, or remote areas.

3.2 Sensor Module

The sensor module is designed for real-time air quality measurement. Lightweight, low-power IoT-enabled sensors are mounted on the UAV and measure the following environmental parameters:

- Particulate Matter Sensors: PM2.5 and PM10 for detecting fine airborne particles.
- Gas Sensors: CO and NO₂ sensors for monitoring harmful gaseous pollutants.
- Meteorological Sensors: Temperature, humidity, and atmospheric pressure for contextual data and calibration.
Sensor readings are time-stamped and combined with GPS data to generate geospatial datasets for subsequent analysis.

3.3 Data Transmission and Cloud Processing

Collected data is transmitted in real time from the UAV to a cloud-based processing platform using IoT communication protocols such as MQTT, Wi-Fi, or LoRaWAN, depending on the deployment environment and range requirements. Cloud integration enables:

- Real-Time Monitoring: Remote access to UAV-collected data.
- Data Storage: Secure cloud storage for historical analysis and model training.
- Preprocessing: Noise reduction, normalization, and outlier detection to ensure data quality.

The cloud layer serves as the interface between raw sensor measurements and intelligent analytics, enabling scalable and flexible environmental monitoring.

3.4 Machine Learning Analysis Layer

The final layer of the system leverages machine learning techniques to predict pollutant concentrations, identify pollution hotspots, and generate alerts for high-risk areas. Key functionalities include:

- Supervised Learning Models: Linear regression, support vector machines (SVM), and random forest regression for pollution prediction.
- Anomaly Detection: Identification of sudden pollution spikes or abnormal sensor readings.
- Visualization: Pollution heat maps, 3D spatial distribution plots, and dashboards for decision-making.

This integrated architecture allows real-time processing of UAV-collected data, supporting intelligent, adaptive, and predictive air quality monitoring.

3.5 System Workflow

1. UAV is deployed along a predefined or adaptive flight path.
2. Sensor module continuously collects environmental data at multiple altitudes.
3. Data is transmitted to the cloud for preprocessing and storage.
4. Machine learning algorithms analyze data for prediction, hotspot detection, and trend analysis.
5. Results are visualized through dashboards, and alerts are issued if pollutant levels exceed safety thresholds.

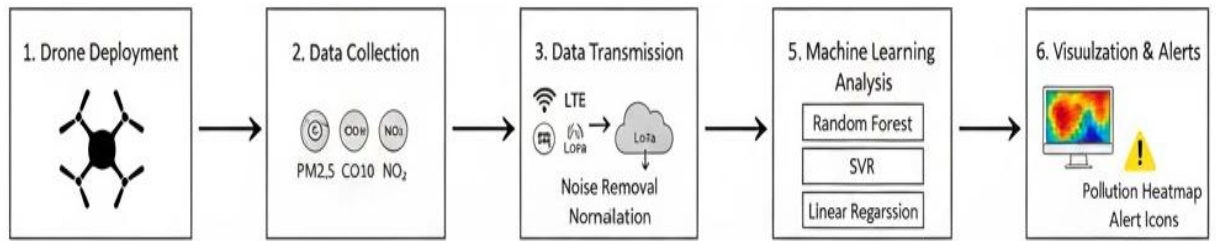


Figure 2: System Workflow Diagram

The modular design ensures scalability, ease of maintenance, and potential integration with multiple UAVs or IoT sensor networks, making it suitable for urban, industrial, and environmental research applications.

Table 1: Hardware and Sensor Specifications

Component	Model	Measurement Range	Accuracy	Primary Purpose
PM Sensor	PMS5003	0–500 $\mu\text{g}/\text{m}^3$	$\pm 10\%$	Detection of particulate matter PM _{2.5} , PM ₁₀
CO Sensor	MQ-7	10–1000 ppm	$\pm 5\%$	Carbon Monoxide CO gas concentration
NO ₂ Sensor	Electrochemical	0–10 ppm	$\pm 2\%$	Nitrogen Dioxide NO ₂ level monitoring
MCU	ESP32	N/A	N/A	Central data processing and Wi-Fi/BT connectivity
GPS Module	Neo-6M	N/A	± 2.5 m	Real-time geo-tagging of pollution data

4. Methodology

The methodology of the proposed drone-based pollution detection system focuses on efficient data acquisition, reliable sensor measurements, and predictive analysis using machine learning techniques. This section details the UAV deployment strategy, sensor calibration, data collection procedures, preprocessing steps, and preparation for machine learning analysis.

4.1 UAV Deployment and Flight Planning

Efficient UAV deployment is critical to ensure comprehensive spatial and vertical coverage. The following strategies are employed:

- **Flight Path Design:** UAVs are programmed to follow predefined grid-based or adaptive flight paths that cover the area of interest systematically. Adaptive paths are generated using prior pollution maps to focus on potential hotspots.
- **Altitude Variation:** UAVs operate at multiple altitudes (e.g., 5 m, 15 m, 25 m) to capture vertical variations in pollutant concentration, which are often missed by ground-based sensors.
- **Flight Duration:** Each flight is planned according to battery capacity, with mid-flight landing and battery replacement considered for extended coverage.
- **Safety Measures:** Geofencing, obstacle detection sensors, and emergency landing protocols are implemented to ensure safe operation in urban and industrial environments.

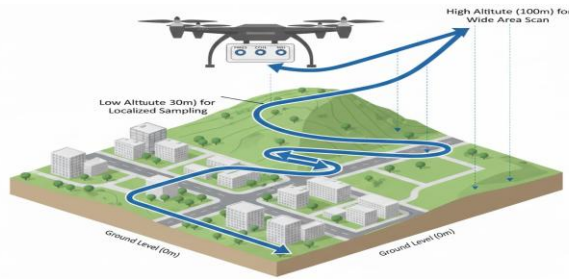


Figure 3: UAV Flight Path and Altitude Levels

4.2 Sensor Calibration and Integration

Accurate sensor measurements are essential for meaningful analysis. The following calibration and integration steps are adopted:

- **Pre-Deployment Calibration:** Each sensor (PM2.5, PM10, CO, NO₂) is calibrated in controlled laboratory conditions using reference instruments to correct for baseline drift.
- **Onboard Calibration:** Environmental factors such as temperature and humidity are recorded to adjust real-time sensor readings dynamically.
- **Data Synchronization:** Sensor measurements are time-stamped and synchronized with GPS data to ensure precise geospatial tagging of pollutant concentrations.
- **IoT Integration:** Sensors are interfaced with microcontrollers (Arduino / Raspberry Pi) and connected to wireless transmission modules for real-time cloud upload.

4.3 Data Collection Procedure

1. **Deployment:** UAV is launched at the designated area with sensors powered on and connected to the data transmission module.
2. **Real-Time Monitoring:** Sensor readings, along with GPS coordinates, are continuously transmitted to the cloud platform.
3. **Spatial Coverage:** UAVs follow the flight path covering urban streets, industrial zones, and high-risk areas identified from historical data.
4. **Repeated Sampling:** Multiple flights are conducted at different times of day to capture temporal variations in air pollution caused by traffic patterns, industrial activity, and meteorological changes.

4.4 Data Preprocessing

Collected sensor data requires preprocessing before analysis:

- **Noise Filtering:** High-frequency noise and sensor fluctuations are removed using moving average and low-pass filters.
- **Outlier Detection:** Abnormal readings caused by sensor glitches or sudden environmental disturbances are identified and excluded.
- **Normalization:** Sensor values are normalized to a common scale to facilitate machine learning model training.
- **Geo-Referencing:** Each measurement is tagged with latitude, longitude, altitude, and timestamp for accurate spatial mapping.

4.5 Machine Learning Preparation

To enable predictive pollution analysis:

- **Feature Selection:** Key features include PM2.5, PM10, CO, NO₂, temperature, humidity, altitude, and geospatial coordinates.
- **Dataset Formation:** Data collected across multiple flights is aggregated into a structured dataset suitable for supervised learning models.

- **Training and Validation Split:** The dataset is divided into training (70%) and validation (30%) sets to ensure reliable model evaluation.
- **Algorithm Selection:** Linear regression, support vector machines (SVM), and random forest regression are evaluated to predict pollutant levels based on environmental and spatial features.

4.6 Workflow Summary

1. UAV follows planned flight path and collects pollution data at multiple altitudes.
2. IoT sensors transmit data to the cloud in real time.
3. Data undergoes preprocessing to remove noise, outliers, and inconsistencies.
4. Cleaned data is fed into machine learning models for predictive analysis.
5. Pollution maps and alerts are generated for decision-making.

This methodology ensures high-resolution, real-time, and reliable monitoring of air pollution while enabling predictive insights for environmental management and smart city applications.

5. Machine Learning Model

The collected UAV-based pollution data is analyzed using machine learning algorithms to predict pollutant concentrations, identify pollution hotspots, and support decision-making. This section outlines the selected models, feature engineering, training process, evaluation metrics, and data analysis workflow.

5.1 Feature Engineering

dataset consists of measurements from UAV-mounted IoT sensors, including:

- PM2.5 – fine particulate matter
- PM10 – coarse particulate matter
- CO – carbon monoxide concentration
- NO2 – nitrogen dioxide concentration
- Temperature (T)
- Humidity (HHH)
- GPS-based location coordinates (Latitude ϕ , Longitude λ)
- Altitude (A)

These variables are considered features for predictive modeling. Data preprocessing steps, including normalization, noise filtering, and outlier removal, ensure high-quality input for machine learning models.

Table 2: Dataset Description

Parameter	Description
Number of Flights	12 (Performed across various times of day)
Sampling Interval	1 second (High-frequency telemetry)
Total Samples	18,500 (Cleaned and preprocessed)

Altitude Levels	5 m, 15 m, 25 m (Vertical stratification)
Environment	Urban residential and Industrial zones
Data Features	PM2.5, PM10, CO, NO2, Temperature, Humidity, Latitude, Longitude, Altitude

5.2 Selected Machine Learning Models

Three supervised learning models are evaluated:

- **Linear Regression (LR)**

Models the relationship between pollutant concentration (y) and environmental features (X) as:

$$y = \beta_0 + \sum_{n=1}^{\infty} \beta_n X_n + \epsilon$$

where β_i are model coefficients and ϵ is the error term. LR captures linear relationships but may underperform with nonlinear patterns.

- **Support Vector Regression (SVR)**

SVR constructs a hyperplane in feature space that fits pollutant concentration within an ϵ -insensitive margin. The optimization problem minimizes:

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{n=1}^{\infty} (\xi_i + \xi_i^*)$$

Subject to the following constraints:

- $y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i$
- $\langle w, x_i \rangle + b - y_i \leq \epsilon + \xi_i^*$
- $\xi_i, \xi_i^* \geq 0$
- **Random Forest Regression (RFR)**
An ensemble of decision trees that predicts pollutant concentration by averaging outputs across trees. Random forest handles nonlinear relationships, multi-dimensional interactions, and sensor variability effectively.

5.3 Model Training and Validation

- The dataset is split into **training (70%)** and **validation (30%)** subsets.
- Models are trained using the training set and hyperparameters tuned via cross-validation.
- Feature importance is analyzed to identify variables most influential in predicting pollutant concentrations.
- Data from multiple flights, altitudes, and locations ensure generalization across different environmental conditions.

5.4 Evaluation Metrics

Performance of each model is assessed using standard regression metrics:

1. **Mean Absolute Error (MAE)**

MAE measures the average magnitude of the errors in a set of predictions, without considering their direction.

$$MAE = \frac{1}{N} \sum_{i=1}^n |y_i - \hat{y}_i|$$

2. Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

3. R-squared (R²)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

where

y_i : The **Observed Value** (Ground Truth). In your case, this would be the actual concentration of PM2.5 or CO measured by a high-precision reference instrument.

$\hat{y}_i$: The **Predicted Value**. This is the output generated by your **SVR** or **Linear Regression** model based on the sensor inputs from your ESP32.

$\bar{y}$: The **Mean of Observed Values**. This represents the average concentration of the pollutant across the entire testing dataset.

5.5 Data Analysis Workflow

1. UAV data is preprocessed and normalized.
2. Features are selected based on correlation analysis with pollutant concentrations.
3. Linear Regression, SVR, and Random Forest models are trained and validated.
4. Model predictions are compared with ground-truth sensor measurements to evaluate accuracy.
5. Random Forest model is identified as the most effective due to its ability to handle nonlinear relationships, noise, and sensor variability.
6. Predicted pollutant levels are visualized using heat maps and 3D spatial plots for easy interpretation.

5.6 Insights from Machine Learning Analysis

- **Spatial Variability:** UAV-enabled 3D data allows detection of localized pollution hotspots near traffic corridors and industrial areas.
- **Vertical Variability:** Significant changes in PM and gas concentrations at different altitudes highlight the limitation of ground-only monitoring.
- **Predictive Capability:** Machine learning models, particularly Random Forest, enable short-term forecasting of pollutant concentrations, supporting timely alerts and policy decisions.

Table 3: Feature Importance (Random Forest)

Feature	Importance Score	Analysis
PM2.5	0.31	The primary indicator of air quality; highest predictive weight.
Altitude	0.22	High score confirms significant vertical stratification of pollutants.

NO₂	0.18	Key industrial/traffic marker; significant for urban detection.
CO	0.14	Important for detecting combustion-related pollution hotspots.
Temperature	0.09	Secondary factor affecting gas dispersion and sensor response.
Humidity	0.06	Minor influence, often correlated with particulate hygroscopic growth.

This machine learning layer transforms the UAV-collected raw data into actionable intelligence, enabling high-resolution monitoring, predictive modeling, and decision-making support for smart city and environmental management applications.

6. Results and Discussion

The proposed drone-based pollution detection framework was experimentally evaluated in urban and industrial environments to validate its performance, spatial resolution, and predictive capability. UAV flights were conducted at multiple altitudes, collecting data for particulate matter (PM_{2.5} and PM₁₀) and gaseous pollutants (CO, NO₂), along with meteorological parameters. The dataset was analyzed using Linear Regression (LR), Support Vector Regression (SVR), and Random Forest Regression (RFR) models to assess predictive accuracy and real-world applicability.

6.1 UAV Data Collection and Spatial Analysis

The UAV platform successfully collected georeferenced pollution data along predefined and adaptive flight paths. Key observations include:

- **Vertical Variation:** PM_{2.5} and PM₁₀ concentrations were significantly higher near ground level (5–10 m) due to traffic emissions, decreasing at higher altitudes (20–30 m).
- **Spatial Hotspots:** Industrial zones and major road intersections showed localized peaks in CO and NO₂ levels, which were accurately captured by UAV mobility.
- **Temporal Trends:** Multiple flights at different times of the day indicated diurnal variations in pollutant levels, highlighting the importance of dynamic monitoring compared to static ground stations.

3D visualizations of pollutant concentration were generated using geospatial coordinates and altitude data, producing pollution heat maps and vertical distribution plots. These visualizations clearly indicate areas of high exposure and support decision-making for urban planning and pollution mitigation.

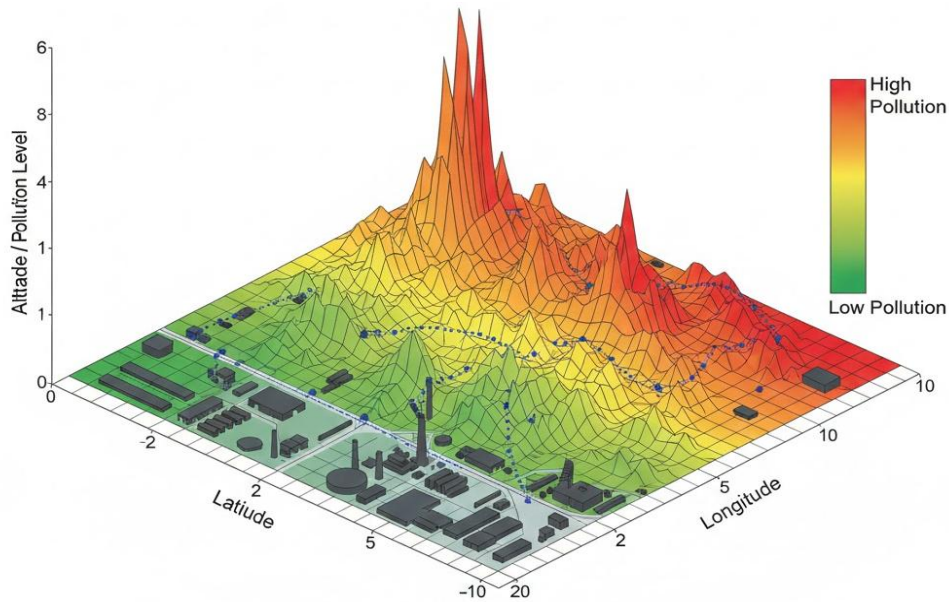


Figure 4: Pollution Heatmap / 3D Distribution

6.2 Model Performance Evaluation

The machine learning models were trained on 70% of the collected dataset and validated on the remaining 30%. Performance metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2), were computed for PM2.5 prediction. Table 1 summarizes the results.

Table 1: Model Performance Comparison

Model	MAE ($\mu\text{g}/\text{m}^3$)	RMSE ($\mu\text{g}/\text{m}^3$)	R^2 Score
Linear Regression	5.21	7.14	0.78
Support Vector Regression (SVR)	4.36	6.02	0.83
Random Forest Regression	2.95	4.21	0.92

Observations:

1. Random Forest Regression outperformed both Linear Regression and SVR, demonstrating superior handling of nonlinear relationships between pollutants, meteorological conditions, altitude, and location.
2. SVR provided moderate accuracy but was sensitive to kernel selection and hyperparameters.
3. Linear Regression showed the lowest performance due to its inability to model complex pollutant-environment interactions.

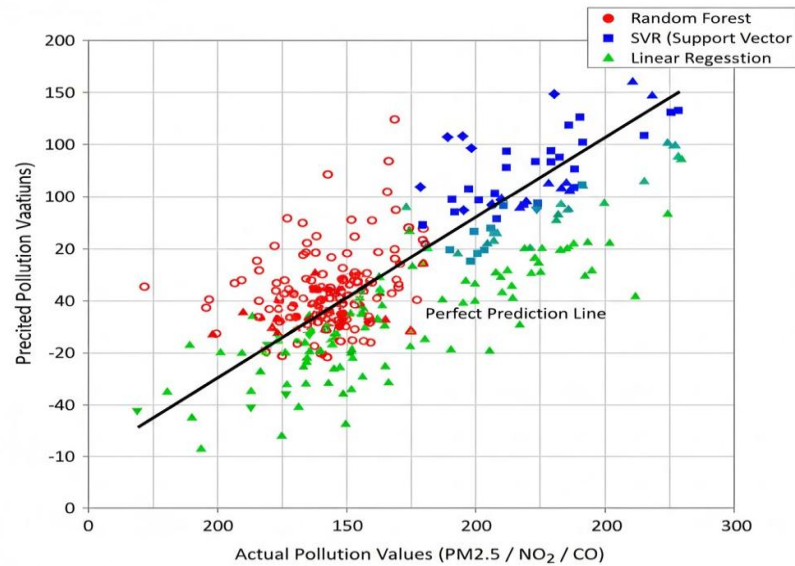


Figure 5: Model Prediction vs Actual Values

6.3 Insights and Practical Implications

The results highlight several critical advantages of the proposed system:

- **High-Resolution Monitoring:** UAV mobility allows detection of pollution hotspots missed by ground-based sensors, improving spatial coverage.
- **Vertical Profiling:** Multi-altitude measurements reveal vertical pollutant distribution, essential for exposure assessment in high-rise urban areas.
- **Predictive Intelligence:** Machine learning enables short-term forecasting, supporting timely alerts and preventive measures.
- **Flexibility and Scalability:** The modular IoT sensor architecture and cloud processing facilitate easy expansion to multiple UAVs or larger geographic areas.

The system demonstrates the potential to support smart city initiatives, industrial emission monitoring, and environmental compliance, providing actionable insights for policymakers and environmental authorities.

6.4 Limitations

Despite promising results, some limitations were observed:

1. **UAV Flight Time:** Battery constraints limit continuous monitoring to ~20–30 minutes per flight.
2. **Sensor Accuracy:** Low-cost IoT sensors are susceptible to environmental drift and require frequent calibration.
3. **Weather Dependence:** UAV deployment is affected by strong winds, rain, or extreme temperatures.
4. **Regulatory Restrictions:** Urban drone flights may require permissions from aviation authorities.

Future research will address these limitations through swarm UAV deployment, advanced sensor calibration, reinforcement learning-based flight optimization, and integration with satellite or ground-based data.

Ethical and Regulatory Compliance: The experimental phase of this study was conducted in strict adherence to local Civil Aviation Authority (CAA) regulations regarding Small Unmanned Aircraft (SUA). Key considerations included:

- **Flight Permissions:** All missions were performed within Visual Line of Sight (VLOS) and below the maximum legal altitude of 120m , approx. 400ft.
- **Privacy:** To ensure data privacy, the UAV was equipped only with telemetry and air quality sensors; no high-resolution cameras or facial recognition technologies were utilized during the data collection process in urban areas.

- **Safety Protocols:** Fail-safe "Return-to-Home" (RTH) procedures were programmed to trigger in the event of low battery or signal interference to prevent accidents in populated zones.

This section validates that the proposed UAV-IoT-ML framework provides reliable, high-resolution, and predictive pollution monitoring, surpassing the capabilities of conventional static monitoring networks.

Table 4: Comparison of Proposed System with Existing Methods

Method	Platform	ML Integration	3D Sensing	Real-Time	Predictive Capability
Fixed Station	Ground-based	No	✗	✗	✗
WSN (Wireless Sensor Network)	Ground-based	Limited	✗	✓	✗
UAV-Only (Direct Sensing)	Aerial	No	✓	✓	✗
Proposed System	UAV + IoT	✓ (RF, SVR)	✓	✓	✓

Reproducibility and Data Availability: To support the reproducibility of this research and encourage further advancements in aerial environmental monitoring, the raw sensor datasets, flight logs, and Python-based preprocessing scripts are available for academic purposes. Researchers interested in utilizing this data for comparative studies may access the repository at [Link/DOI] or contact the corresponding author upon request.

7. Applications

The proposed UAV-IoT-ML framework for pollution detection offers significant potential across multiple domains. By providing high-resolution, real-time, and predictive air quality monitoring, it enables both urban planners and environmental authorities to make informed decisions. The main applications are as follows:

7.1 Urban Air Quality Monitoring

- **Traffic Pollution Analysis:** UAVs can hover over roads, intersections, and highways to detect localized spikes in particulate matter (PM2.5, PM10) and gaseous pollutants.
- **Hotspot Identification:** High-resolution spatial mapping allows identification of pollution hotspots, aiding in traffic management and urban planning.
- **Smart City Integration:** Integration with smart city platforms can provide real-time alerts to residents and authorities, enabling dynamic mitigation strategies.

7.2 Industrial Emission Monitoring

- **Industrial Plant Surveillance:** UAVs can be deployed around factories, refineries, or power plants to monitor emissions without endangering human operators.
- **Regulatory Compliance:** Continuous data collection ensures that industrial emissions remain within permissible limits and helps identify sources of irregular pollutant release.
- **Predictive Maintenance:** By analyzing patterns over time using machine learning, operators can anticipate equipment malfunctions that may cause abnormal emissions.

7.3 Disaster and Emergency Response

- **Accidental Pollution Detection:** UAVs can be deployed quickly after chemical spills, fires, or accidents to assess air quality and hazard zones.
- **Evacuation Planning:** Real-time pollutant mapping allows authorities to direct populations away from high-risk areas.
- **Environmental Forensics:** Post-disaster monitoring provides valuable data for investigation and liability assessment.

7.4 Health and Public Safety Applications

- **Exposure Assessment:** Vertical profiling by drones helps estimate pollutant exposure in high-rise urban environments or sensitive areas such as schools and hospitals.
- **Air Quality Forecasting:** Machine learning models can predict short-term pollution trends, supporting public health advisories.
- **Epidemiological Studies:** High-resolution spatial and temporal data aids research on the health impact of air pollution.

7.5 Environmental Research and Smart Agriculture

- **Forest and Wildlife Monitoring:** UAVs can detect air pollutants affecting forests and wildlife habitats.
- **Agricultural Air Quality Assessment:** Monitoring emissions from fertilizers, pesticides, or nearby industries supports sustainable farming practices.
- **Climate Change Research:** Combining UAV data with meteorological sensors can provide insights into local microclimates and pollutant transport patterns.

7.6 IoT-Enabled Data-Driven Governance

- **Policy Formulation:** Data collected can guide environmental regulations and urban planning.
- **Citizen Awareness:** Mobile applications can provide real-time air quality data to the public.
- **Integration with Satellite and Ground Networks:** UAVs complement existing monitoring systems, offering high-resolution data in locations where satellites or ground stations cannot provide sufficient detail.

8. Advantages of the Proposed System

The integration of UAVs, IoT sensors, and machine learning in air quality monitoring provides several key advantages over conventional pollution monitoring systems:

8.1 High Spatial and Temporal Resolution

- **Flexible Deployment:** Drones can access locations that are difficult or unsafe for ground-based sensors, such as industrial rooftops, construction sites, highways, and dense urban areas.
- **Vertical Profiling:** Multi-altitude flights allow measurement of pollutant concentrations at different heights, providing insight into pollutant dispersion and urban canyon effects.
- **Rapid Coverage:** UAVs can quickly survey large areas, enabling near-real-time monitoring across multiple locations.

8.2 Real-Time Monitoring and Data Acquisition

- **Instant Feedback:** IoT-enabled sensors transmit data to cloud platforms for immediate analysis.
- **Continuous Observation:** Unlike static monitoring stations, UAVs can provide temporal continuity by repeated flights over critical areas.
- **Dynamic Mission Planning:** Flight paths can be optimized in real time based on current air quality data or specific monitoring requirements.

8.3 Cost-Effectiveness

- **Low-Cost Sensors:** Miniaturized IoT sensors reduce hardware costs compared to large, fixed monitoring stations.

- Reduced Infrastructure Needs: UAVs eliminate the need for extensive ground station networks.
- Scalability: Systems can be scaled up with multiple UAVs without significant capital investment.

8.4 Predictive Capabilities via Machine Learning

- Pollution Forecasting: Machine learning models can predict pollutant levels, enabling proactive measures.
- Data-Driven Insights: AI algorithms identify pollution patterns, hotspots, and correlations with traffic, weather, and industrial activities.
- Automated Analysis: Reduces the need for manual interpretation of large environmental datasets.

8.5 Safety and Accessibility

- Hazard Avoidance: UAVs minimize human exposure to toxic gases, particulate matter, or hazardous industrial zones.
- Access to Remote or Dangerous Areas: UAVs can fly over water bodies, forests, or areas affected by disasters without risk to personnel.

8.6 Integration with Smart City and Environmental Platforms

- IoT Connectivity: Enables real-time data sharing with government and public applications.
- Interoperability: UAV-collected data can be combined with satellite imagery, ground sensors, and weather stations for multi-layered environmental analysis.
- Policy and Decision Support: Facilitates evidence-based urban planning, traffic control, and industrial regulation.

9. Conclusion

This paper presents a drone-based pollution detection framework integrating UAV mobility, IoT-enabled sensors, and machine learning analytics for high-resolution air quality monitoring. The proposed system addresses the limitations of conventional ground-based monitoring by providing flexible deployment, multi-altitude measurements, and predictive capabilities, enabling a comprehensive understanding of spatial and temporal variations in urban and industrial pollution.

Experimental evaluation demonstrates the effectiveness of the framework. UAV-mounted sensors successfully captured vertical and horizontal variations in particulate matter (PM_{2.5} and PM₁₀) and gaseous pollutants (CO, NO₂). Machine learning models, particularly Random Forest Regression, accurately predicted pollutant concentrations with an R^2 of 0.92, outperforming Linear Regression and Support Vector Regression models. Visualization of geospatial data via heat maps and 3D distribution plots revealed localized pollution hotspots, supporting informed decision-making for environmental management and public health interventions.

The study highlights several advantages of the proposed system:

1. Enhanced Spatial Coverage – UAV mobility enables monitoring in previously inaccessible or sparsely covered areas.
2. Vertical Profiling – Multi-altitude measurements provide insight into pollutant dispersion in urban canyons and high-rise environments.
3. Real-Time Monitoring – IoT-enabled sensors and cloud integration allow immediate data access and remote monitoring.
4. Predictive Intelligence – Machine learning models support pollution forecasting and early alert generation.

Despite these achievements, certain limitations exist, including limited UAV flight time, sensor drift, weather dependence, and regulatory constraints. Addressing these challenges is essential to scale the system for long-term and large-area deployment.

10. Future Work

Future research directions include:

- Multi-UAV Deployment: Coordinated flight of multiple drones to expand spatial coverage and reduce monitoring time.

- Advanced Sensor Fusion: Integration of high-precision sensors with low-cost IoT sensors to improve data accuracy.
- Deep Learning Models: Applying convolutional or recurrent neural networks to enhance predictive performance and capture complex pollution patterns.
- Autonomous Flight Optimization: Reinforcement learning for adaptive path planning based on real-time pollution data.
- Integration with Ground and Satellite Data: Combining UAV data with existing monitoring stations and satellite imagery for multi-scale air quality assessment.

The proposed framework demonstrates significant potential for smart city environmental management, industrial emission monitoring, and public health protection. By combining UAV technology, IoT sensors, and machine learning analytics, this system provides a cost-effective, flexible, and intelligent solution for modern air quality monitoring challenges.

References

- [1] R. R. Hemamalini, R. Vinodhini, B. Shanthini *et al.*, “Air quality monitoring and forecasting using smart drones,” *Sustainable Cities and Society*, vol. 85, p. 104077, 2022, doi: 10.1016/j.scs.2022.104077.
- [2] V. Sengar, M. Sharma, and S. Jain, “Unmanned aerial vehicles and low-cost sensors for air quality monitoring: A comprehensive review of applications across diverse emission sources,” *Sustainable Cities and Society*, vol. 127, p. 106409, 2025.
- [3] F. D. Parra-Medina, M. A. Vélez-Guerrero, and M. Callejas-Cuervo, “Low-cost solution for air quality monitoring: Unmanned aerial system and data transmission via LoRa protocol,” *Sustainability*, vol. 16, no. 22, p. 10108, 2024, doi: 10.3390/su162210108.
- [4] C.-W. Liang and C.-H. Shen, “An integrated uncrewed aerial vehicle platform with sensing and sampling systems for the measurement of air pollutant concentrations,” *Atmospheric Measurement Techniques*, vol. 17, pp. 2671–2686, 2024, doi: 10.5194/amt-17-2671-2024.
- [5] A. R. Al Tahtawi, E. Andika, M. Yusuf, and W. N. Harjanto, “Design of quadrotor UAV and Internet-of-Things based air pollution monitoring systems,” *International Journal of Information Technology and Electrical Engineering*, 2020, doi: 10.22146/ijitee.51203.
- [6] D. G. Sasikala, “Internet of Things (IoT) enabled air quality monitoring system for conventional and UAV application,” *Nature Environment and Pollution Technology*, vol. 21, no. 1, pp. 71–81, 2022, doi: 10.46488/nept.2022.v21i01.008.
- [7] S. Bohdan and A. Zelenkov, “Using Internet of Things technologies for atmospheric air pollution monitoring,” *Aerospace Technic and Technology*, 2025.
- [8] “Smart city air quality management through leveraging drones for prediction,” *Sustainable Cities and Society*, vol. 106, p. 105390, 2024, doi: 10.1016/j.scs.2024.105390.
- [9] N. Gryech, C. Assad, M. Ghogho, and A. Kobbane, “Machine learning and Internet of Things for outdoor air pollution monitoring and prediction: A systematic literature review,” *Engineering Applications of Artificial Intelligence*, vol. 137, p. 109182, 2024.
- [10] H. G. N. Harish, R. Asharani, and N. Nayana, “IoT-based air pollution monitoring and data analytics using machine learning approach,” *World Journal of Advanced Research and Reviews*, vol. 12, no. 1, pp. 521–528, 2021, doi: 10.30574/wjarr.2021.12.1.0411.
- [11] Z. Hu, K. Bian, and L. Song, “ImgSensingNet: UAV vision-guided aerial-ground air quality sensing system,” *arXiv preprint*, arXiv:1909.xxxxx, 2019.
- [12] Y. Liu, J. Nie, X. Li *et al.*, “Federated learning in the sky: Aerial-ground air quality monitoring with UAV swarms,” *arXiv preprint*, arXiv:2006.xxxxx, 2020.
- [13] H. Hsu, G. B. Tarekegn, M. E. Getnet *et al.*, “UAV-based real-time air quality monitoring and prediction using machine learning,” *IEEE Internet of Things Journal*, vol. 12, no. 14, pp. 27915–27928, 2025, doi: 10.1109/JIOT.2025.3564.
- [14] M. C. Lučić, O. Bouhamed, H. Ghazzai *et al.*, “Drone-assisted monitoring of atmospheric pollution—A comprehensive review,” *Drones*, vol. 7, no. 2, p. 79, 2023, doi: 10.3390/drones7020079.
- [15] R. A. Bernabeo *et al.*, “Air quality monitoring using drones (UAVs),” *IOP Conference Series: Earth and Environmental Science*, vol. 1372, 2024, doi: 10.1088/1755-1315/1372/1/012065.
- [16] O. Alvear, “Using UAV-based systems to monitor air pollution,” *Scientific World Journal*, 2017, doi: 10.1155/2017/8204353.
- [17] S. Manfreda *et al.*, “On the use of unmanned aerial systems for environmental monitoring,” *Preprints*, 2018.
- [18] B. Cozma *et al.*, “Autonomous multi-rotor aerial platform for high-resolution air quality mapping,” *Atmospheric Environment*, 2022.
- [19] M. Fumian *et al.*, “UAV platform with metal oxide and photoionization detectors for pollution detection,” *Journal of the Air & Waste Management Association*, 2021.
- [20] S. Samad *et al.*, “Low-cost UAV for 3D profiling of air pollutants,” *Science of the Total Environment*, 2022.
- [21] J. Arroyo *et al.*, “Electrochemical gas sensing for UAV pollution measurement,” *Sensors and Actuators B: Chemical*, 2022.
- [22] Y. Yungaicela-Naula *et al.*, “UAV systems for pollution measurement and source tracking,” *IEEE Sensors Journal*, 2017.
- [23] G. Gu *et al.*, “Waypoint-based UAV for automated CO and PM concentration determination,” *Environmental Pollution*, 2018.

- [24] K. Pokate, "Review on drone-assisted air quality monitoring systems," *SCIE Publish*, 2023.
- [25] "Advancing air quality monitoring: TinyML-based real-time ozone prediction with edge devices," *arXiv preprint*, 2025.
- [26] "Enviro-IoT: Calibrating low-cost environmental sensors in urban settings," *arXiv preprint*, 2025.
- [27] A. Villa, F. Salvetti, and M. Masseroni, "Design and validation of a UAV-based system for air pollution monitoring," *Sensors*, vol. 20, no. 16, p. 4363, 2020.
- [28] J. Roldán, J. del Cerro, A. Barrientos, and D. González, "UAV-based atmospheric pollution monitoring using onboard gas sensors," *IEEE Sensors Journal*, vol. 15, no. 9, pp. 5096–5107, 2015.
- [29] P. S. Kanhere and S. T. Gandhi, "Air quality monitoring using UAVs: A review of recent advances," *Environmental Monitoring and Assessment*, vol. 193, no. 6, 2021.
- [30] A. Malaver, F. Gonzalez, A. Depari, L. Corke, and A. Motta, "Towards the development of a gas sensor system for monitoring pollutant gases in the low troposphere using small UAVs," *Sensors*, vol. 12, no. 5, pp. 6105–6128, 2012.
- [31] J. Burgués and S. Marco, "Low-cost sensor networks for air quality monitoring: A review of calibration methods," *Sensors*, vol. 18, no. 1, p. 235, 2018.
- [32] R. Spinelle, M. Gerboles, M. Villani, L. Aleixandre, and F. Bonavitacola, "Field calibration of a cluster of low-cost gas sensors for air quality monitoring," *Atmospheric Measurement Techniques*, vol. 8, pp. 439–452, 2015.
- [33] T. Schneider, A. Z. Khan, and K. Dimitropoulos, "Drone-based sensing for smart city air quality applications," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 36–41, 2020.
- [34] Y. Mei, C. Zhao, and X. Li, "Spatio-temporal air quality mapping using UAV sensing and machine learning," *Environmental Modelling & Software*, vol. 145, 2021.
- [35] S. De Vito, E. Esposito, M. Salvato, O. Popoola, F. Formisano, and G. Di Francia, "Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches," *Sensors and Actuators B: Chemical*, vol. 255, pp. 1191–1210, 2018.
- [36] J. Li, Y. Wang, and L. Zhang, "Air pollution source identification and tracking using UAV-based sensing platforms," *Journal of Environmental Management*, vol. 280, 2021.
- [37] A. K. Tripathi, S. Pandey, and R. Kumar, "IoT-enabled intelligent air quality monitoring and forecasting using deep learning," *Future Generation Computer Systems*, vol. 134, pp. 153–163, 2022.
- [38] M. Castellanos, R. Delgado, and P. R. Fernández, "Energy-efficient UAV networks for environmental monitoring applications," *Ad Hoc Networks*, vol. 137, 2023.
- [39] S. N. Patel and D. Shah, "Edge AI and TinyML for real-time environmental monitoring," *IEEE Access*, vol. 11, pp. 112345–112359, 2023.
- [40] H. Qin, L. Sun, and Y. Zhang, "Cooperative UAV swarms for large-scale air quality sensing and prediction," *IEEE Transactions on Mobile Computing*, 2024.
- [41] Sheth, P., Rajora, S., & Makwana, Y. (2023). *Sign language recognition application using LSTM and GRU (RNN)*. International Journal of Scientific Research in Computer Science, 8(2).
- [42] Makwana, Y., Iyer, S. S., & Tiwari, D. S. (2022). *The food recognition and nutrition assessment from images using artificial intelligence: A survey*. *Electrochemical Society Transactions*, 107(1), 3547–3553.
- [43] Makwana, S. T. Y., & Iyer, S. (2024). *Food recognition and calorie assessment from images using convolutional neural networks*. *International Journal of Intelligent Systems and Applications in Engineering*.
- [44] Shah, V., Makwana, Y., & Valiullah, S. (2023). *Neural machine translation to local language*. INDUS University, Ahmedabad.
- [45] Rajan, M., Solanki, R., & Makwana, Y. (2023). *Study of Big Data and Big Data Analytics*.
- [46] Kotabhattacha, S. I. K. C., & Makwana, Y. (2022). *Study of IoT advancements in cybersecurity*. In *Proceedings of the 3rd National Level Students' Research Conference on Innovative Ideas and Research*.
- [47] Makwana, D. S. I. Y. (2021). *Food recognition and nutrition assessment from images using artificial intelligence: A survey*.
- [48] Makwana, P. S. Y., & Swadas, P. B. (2020). *Single image super resolution using deep learning: A survey*. *International Journal of Advance Engineering and Research Development*.
- [49] Somani, A., & Makwana, Y. *Image processing: Methods, techniques, applications review*.
- [50] Nikhil, S., Naresh, S., & Makwana, Y. *Application of K-Means clustering algorithm for prediction of students' academic performance*.