DOI: 10.21917/ijsc.2025.0552

TRANSFORMER-DRIVEN DEEP LEARNING IOT-BASED COGNITIVE RADIO-BASED SENSOR NETWORKS IN REAL-TIME AIR POLLUTION MONITORING

Mayur N. Bhalia[1] and G.P. Suja[2]
[1]*Department of Electronics and Communication Engineering, Government Polytechnic, Rajkot, India*
[2]*Department of Computer Science, Muslim Arts College, India*

Abstract
Rapid urbanization and industrialization have significantly increased air pollution levels, adversely affecting public health and environmental sustainability. Traditional monitoring systems often suffer from limited spatial coverage and delayed data analysis, making real-time pollution assessment challenging. Existing approaches struggle to efficiently process the large-scale, heterogeneous data generated by IoT-enabled Cognitive Radio-Based Sensor Networks (CRSNs). Conventional machine learning models often fail to capture complex temporal and spatial patterns in pollution dynamics, limiting predictive accuracy and early warning capabilities. This study proposes a transformer-based deep learning framework combined with IoT-enabled CRSNs for accurate and real-time air pollution monitoring. The CRSN comprises distributed sensors collecting continuous data on particulate matter (PM2.5, PM10), NO_2, CO, O_3, and other pollutants. The transformer model controls its self-attention mechanism to capture temporal dependencies and inter-sensor correlations, enabling robust prediction of pollution trends. Data preprocessing involves normalization, anomaly detection, and feature embedding to enhance model performance. Comparative experiments are conducted against conventional LSTM and GRU models to evaluate prediction accuracy and system responsiveness. Experimental results establish that the transformer-based model achieves superior performance with a mean absolute error (MAE) of 4.2 $\mu g/m^3$ for PM2.5 prediction, outperforming LSTM (MAE = 6.1 $\mu g/m^3$) and GRU (MAE = 5.8 $\mu g/m^3$). The proposed system provides accurate, fine-grained pollution maps in real-time, enabling timely alerts and informed decision-making for environmental authorities.

Keywords:
IoT, Cognitive Radio-Based Sensor Networks, Air Pollution Monitoring, Transformer Deep Learning, Real-Time Prediction

1. INTRODUCTION

Air pollution has emerged as one of the most pressing environmental and public health challenges in the modern era. Rapid industrialization, urbanization, and increased vehicular emissions have led to elevated concentrations of pollutants such as particulate matter (PM2.5 and PM10), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3), significantly affecting respiratory and cardiovascular health [1]. Monitoring air quality is critical not only for regulatory compliance but also for enabling informed decisions in urban planning, environmental protection, and public safety. Recent advances in Internet of Things (IoT) technology have enabled the deployment of dense Cognitive Radio-Based Sensor Networks (CRSNs) capable of continuously collecting real-time environmental data across urban landscapes [2]. These networks provide a promising solution for dynamic air quality monitoring, overcoming the limitations of traditional, sparsely located monitoring stations that fail to capture localized pollution events [3].

Despite the potential of IoT-enabled CRSNs, several challenges hinder their full utilization. First, the large volume of heterogeneous data generated by distributed sensors presents significant processing and storage demands [4]. Second, sensor measurements are often affected by environmental noise, calibration drift, and intermittent connectivity, which can compromise the reliability and accuracy of collected data [5]. Addressing these issues requires robust computational models capable of handling uncertainty, integrating multi-source data, and providing accurate, timely predictions.

The existing approaches for air pollution prediction face additional problems. Conventional statistical and machine learning models, including regression-based methods and recurrent neural networks, often struggle to capture complex temporal and spatial correlations in pollutant dynamics [6]. Moreover, these models typically require manual feature engineering and are prone to overfitting when dealing with high-dimensional, continuous data streams from CRSNs [7]. As a result, real-time forecasting and fine-grained pollution mapping remain significant challenges, limiting the practical effectiveness of air quality monitoring systems [8].

To address these limitations, this study aims to develop a transformer-based deep learning framework for IoT-enabled CRSNs, enabling accurate, real-time air pollution monitoring. The primary objectives are: (i) to design a scalable and robust architecture capable of handling multi-sensor, high-dimensional data; (ii) to leverage transformer-based self-attention mechanisms to capture both temporal trends and spatial correlations in pollutant concentrations; and (*i*) to evaluate the proposed approach against existing recurrent models such as LSTM and GRU in terms of prediction accuracy, reliability, and computational efficiency.

The novelty of this research lies in integrating transformer-based deep learning with IoT-enabled CRSNs for environmental monitoring—a combination that, to the best of our knowledge, has been minimally explored in urban air quality applications. Unlike conventional models, the transformer can simultaneously model long-range dependencies and inter-sensor relationships without requiring extensive feature engineering, providing both accuracy and interpretability in real-time predictions.

The key contributions of this study are twofold. First, we propose a transformer-driven predictive framework tailored for real-time air quality monitoring, capable of handling noisy, high-volume sensor data with high precision. Second, we demonstrate the system's practical applicability by conducting extensive experiments comparing it with traditional deep learning models, showing improved mean absolute error (MAE), enhanced robustness to missing data, and scalability for deployment in smart city environments. Collectively, these contributions advance the state-of-the-art in IoT-based environmental monitoring and provide actionable insights for policymakers and urban planners seeking to mitigate the adverse effects of air pollution.

2. RELATED WORKS

Over the past decade, significant research efforts have been directed toward leveraging IoT-enabled CRSNs and advanced machine learning techniques for air pollution monitoring and prediction. Traditional approaches relied on statistical models and linear regression techniques to estimate pollutant concentrations; however, these methods often fail to capture the nonlinear, dynamic behavior of urban air pollution [6]. While effective for small-scale datasets, conventional models are limited in scalability and are unable to process the continuous, high-dimensional data streams produced by dense CRSN deployments. These limitations have motivated researchers to explore deep learning-based approaches that can handle complex spatiotemporal patterns inherent in air quality data.

Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, have been widely employed to address temporal dependencies in pollution data [7]. LSTM models have shown promising results in short-term forecasting of PM2.5 and other pollutants, enabling early warning systems for urban populations. However, LSTMs still exhibit challenges when modeling long-range temporal dependencies and multiple interacting sensors simultaneously. Additionally, they often require extensive tuning and preprocessing to handle missing or noisy sensor readings, which is common in real-world IoT deployments. Variants such as gated recurrent units (GRUs) have been proposed to reduce computational complexity while maintaining predictive performance, but their capacity to capture multi-sensor correlations across a city-wide network remains limited [8].

To overcome these challenges, hybrid models that combine spatial and temporal information have gained attention. Convolutional neural networks (CNNs) combined with LSTM layers have been applied to generate spatiotemporal air pollution maps, leveraging CNNs to extract local spatial features and LSTMs to model temporal dynamics [9]. While effective for medium-scale networks, these models often struggle with high-dimensional data and long-term dependencies, limiting their real-time applicability. Moreover, the computational overhead associated with CNN-LSTM architectures can hinder deployment on resource-constrained edge devices in IoT networks.

Recent research has turned toward transformer-based architectures, originally developed for natural language processing, for handling sequential and high-dimensional sensor data [10]. Transformers utilize self-attention mechanisms to weigh the relevance of different input features dynamically, enabling them to model long-range dependencies efficiently. Early studies have demonstrated the potential of transformers in environmental applications, including air pollution prediction and energy consumption forecasting, showing improved accuracy and faster convergence compared to traditional RNN-based models. By capturing both temporal trends and inter-sensor correlations, transformer models provide a scalable solution for dense urban CRSNs.

In addition to predictive modeling, several studies have emphasized the importance of IoT-enabled CRSN design for real-time monitoring. Sensor placement optimization, fault-tolerant network protocols, and energy-efficient data transmission strategies have been explored to enhance data reliability and system longevity [11]. Integrating advanced predictive models with well-designed sensor networks allows not only for accurate forecasting but also for proactive decision-making, such as dynamic traffic management or targeted emission control measures.

Despite these advances, there remain gaps in practical implementation. Most existing studies focus on either model development or sensor network design in isolation, rather than combining both into an end-to-end, real-time monitoring framework [12]. Furthermore, the interpretability of deep learning models in environmental contexts is limited, making it challenging for policymakers to understand the rationale behind predictions. Addressing these gaps requires a comprehensive framework that combines transformer-based modeling with robust IoT sensor networks, providing accurate, interpretable, and actionable insights for urban air quality management.

3. PROPOSED METHOD

The proposed method combines IoT-enabled CRSNs with a transformer-based deep learning model for accurate and real-time air pollution monitoring. The framework begins with a distributed CRSN that continuously collects environmental data, including PM2.5, PM10, NO_2, CO, and O_3 levels, across urban areas. The collected data undergo preprocessing steps such as normalization, missing value imputation, and feature embedding to ensure quality and consistency.

The preprocessed data is then fed into a transformer model, where the self-attention mechanism captures temporal dependencies and inter-sensor correlations, enabling robust prediction of pollution trends. Finally, the model outputs fine-grained real-time air quality predictions, which can be visualized as pollution maps or alerts for decision-making.

3.1 SENSOR DEPLOYMENT AND DATA COLLECTION

The first step in the proposed framework involves the deployment of a dense IoT-enabled CRSN across the target urban environment.

Each sensor node is equipped with pollutant-specific sensing modules capable of measuring PM2.5, PM10, NO_2, CO, O_3, temperature, and humidity in real-time. These nodes communicate via low-power wireless protocols (e.g., LoRaWAN or ZigBee) to a central gateway, which aggregates the data for further processing.

The spatial distribution of sensors is designed to maximize coverage in high-traffic areas and industrial zones, while also ensuring redundancy to handle potential node failures. Temporal resolution is configured to capture rapid fluctuations in pollutant concentrations, typically with measurements taken every 5–10 minutes.

Table.1. Sensor Data Snapshot

Sensor ID	Location	PM2.5 (μg/m³)	PM10 (μg/m³)	NO_2 (ppb)	CO (ppm)	O_3 (ppb)	Time stamp
S01	Downtown	56	82	30	0.8	45	2025-10-14 10:00
S02	Industrial	68	95	42	1.2	38	2025-10-14 10:00
S03	Residential	35	48	18	0.5	50	2025-10-14 10:00

The Table.1 provides a representative snapshot of the raw sensor data collected across different urban zones. These heterogeneous data streams form the foundation for subsequent preprocessing and modeling. The spatial and temporal variability of pollutants can be represented mathematically as:

$$X(t,s) = \left[x_1(t,s), x_2(t,s), \ldots, x_m(t,s)\right] \tag{1}$$

where $X(t,s)$ denotes the pollutant concentration vector at time t and sensor location s, and mmm represents the number of pollutant types. This multi-dimensional data is transmitted to the central processing unit for further cleaning and feature extraction.

3.2 DATA PREPROCESSING

Collected sensor data often contains noise, missing values, or outliers due to environmental interference or sensor malfunction. Therefore, preprocessing is critical to ensure data quality before feeding it into the transformer model. The preprocessing pipeline includes:

- **Missing Value Imputation:** Linear interpolation or k-nearest neighbor (KNN) imputation is used to fill gaps in sensor readings.
- **Normalization:** Min-max normalization is applied to scale pollutant values between 0 and 1 for improved model convergence.
- **Feature Embedding:** Each sensor reading is transformed into a feature vector incorporating pollutant levels, timestamp embeddings, and spatial coordinates.

Table.2. Preprocessed Sensor Data

Sensor ID	PM2.5 (norm)	PM10	NO_2	CO	O_3	Time Embedding	Location Embedding
S01	0.45	0.55	0.33	0.40	0.50	0.10	[0.2, 0.3]
S02	0.55	0.63	0.47	0.60	0.42	0.10	[0.8, 0.1]
S03	0.28	0.32	0.20	0.25	0.55	0.10	[0.1, 0.7]

The Table.2 illustrates the normalized and embedded sensor data, ready for input into the transformer network. The feature embedding can be expressed as:

$$\mathbf{f}_i = [\frac{x_{i1} - x_{min,1}}{x_{max,1} - x_{min,1}}, \ldots, \frac{x_{im} - x_{min,m}}{x_{max,m} - x_{min,m}}] \oplus \mathbf{e}_t \oplus \mathbf{e}_s \tag{2}$$

where $\mathbf{f}_i$ is the embedded feature vector for sensor i, $\mathbf{e}_t$ is the temporal embedding, $\mathbf{e}_s$ is the spatial embedding, and $\oplus$ denotes concatenation.

3.3 TRANSFORMER-BASED MODELING

The preprocessed data is fed into a transformer network, which utilizes self-attention to learn complex temporal dependencies and spatial correlations among sensors. Unlike RNNs, the transformer allows parallel processing of sequences and effectively captures long-range dependencies, improving prediction accuracy. The model architecture consists of an encoder-decoder setup where the encoder processes historical pollutant sequences, and the decoder predicts future concentrations. Multi-head attention layers compute the importance of each sensor reading relative to others, enabling the model to weigh contributions dynamically.

Table.3. Predicted vs Actual PM2.5 Values

Timestamp	Actual PM2.5	Predicted PM2.5	Error (μg/m³)
2025-10-14 10:00	56	54	2
2025-10-14 10:05	58	57	1
2025-10-14 10:10	60	61	1

The Table.3 demonstrates the transformer's predictive performance, showing high accuracy and low error. The self-attention mechanism is mathematically represented as:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right) \tag{3}$$

where Q, K, and V are query, key, and value matrices derived from input features, and d_k is the dimensionality of the key vectors. This equation allows the model to learn interdependencies between all sensors across time efficiently.

3.4 PREDICTION

Once trained, the transformer outputs real-time pollutant forecasts for each sensor node. These predictions are aggregated to generate dynamic air quality maps, providing a spatially resolved view of urban pollution. Alerts are generated for regions exceeding threshold pollutant levels, enabling proactive decision-making for environmental management.

Table.4. Real-Time Air Quality Alert

Location	Predicted PM2.5	AQI Category	Alert Level
Downtown	65	Unhealthy	High
Industrial	72	Unhealthy	High
Residential	40	Moderate	Medium

The Table.4 shows actionable insights derived from transformer predictions, useful for city authorities and public advisories. The final prediction can be formulated as:

$$\hat{X}(t+\Delta t, s) = f_\theta(\mathbf{F}_t) \tag{4}$$

where $\hat{X}(t+\Delta t, s)$ is the forecasted pollutant vector at time $t+\Delta t$ for sensor s, and f_θ represents the trained transformer network applied to the feature sequence $\mathbf{F}_t$.

The model's performance is evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2. Optimization involves hyperparameter tuning of attention heads, embedding dimensions, and learning rates to balance accuracy and computational efficiency. The model is also tested for robustness against missing sensor data and noise.

Table.5. Model Performance Metrics

Model	MAE (µg/m³)	RMSE (µg/m³)	R^2
LSTM	6.1	7.8	0.82
GRU	5.8	7.5	0.84
Transformer	4.2	5.1	0.91

The Table.5 highlights the superior performance of the transformer model over traditional deep learning approaches. Mathematically, MAE and RMSE are expressed as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|, \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{6}$$

where $\hat{y}_i$ and y_i are the predicted and actual pollutant concentrations, respectively.

4. EXPERIMENTS

The experimental evaluation of the proposed transformer-based air pollution monitoring framework was conducted using Python 3.11 with TensorFlow 2.13 as the primary deep learning library. Simulations were performed on a workstation equipped with an Intel Core i9-13900K CPU, 64 GB RAM, and an NVIDIA RTX 4090 GPU to support parallel training of high-dimensional IoT sensor data. The IoT-CRSN environment was emulated using a combination of real-world sensor datasets and synthetic data streams generated to simulate temporal and spatial variations in pollutant concentrations. Sensor readings were collected at a 5-minute interval for PM2.5, PM10, NO_2, CO, and O_3. Preprocessing, feature embedding, and transformer training were performed on this platform, with hyperparameter tuning conducted using grid search and cross-validation to optimize model performance. The experimental setup includes both CRSN parameters and transformer model configurations. The Table.6 summarizes the key parameters used in the simulation.

Table.6. Experimental Setup

Parameter	Value / Setting
Number of sensors	50
Sampling interval	5 min
Pollutants measured	PM2.5, PM10, NO_2, CO, O_3
Transformer encoder layers	4
Transformer decoder layers	4
Attention heads	8
Embedding dimension	128
Batch size	64
Learning rate	0.001
Epochs	100

The Table.6 highlights the critical experimental settings and transformer configurations, which ensure reproducible results and robust performance evaluation.

4.1 PERFORMANCE METRICS

The evaluation of the proposed framework was conducted using metrics:

- **Mean Absolute Error (MAE):** It measures the average magnitude of prediction errors. Lower values indicate better accuracy.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{7}$$

- **Root Mean Square Error (RMSE):** It provides an aggregated measure of error magnitude, penalizing larger deviations.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{8}$$

- **Coefficient of Determination (R^2):** It represents the proportion of variance in the measured data explained by the model.

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

- **Mean Absolute Percentage Error (MAPE):** It indicates relative prediction accuracy in percentage terms.

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad ()$$

- **Prediction Latency (PL):** It measures the computational efficiency of the framework in generating real-time forecasts, crucial for IoT-enabled monitoring.

4.2 DATASET

The evaluation utilized a combination of real-world and synthetic datasets. Real-world data was sourced from urban air quality monitoring stations, comprising pollutant measurements over a period of 12 months. Synthetic data streams were generated to simulate additional sensor nodes for urban expansion scenarios. The Table.7 provides a brief description of the dataset.

Table.7. Dataset Description

Dataset Source	No. of Samples	Features	Duration
Real-World Sensor Data	105,120	PM2.5, PM10, NO_2, CO, O_3	12 months

Synthetic CRSN Simulation	50,000	PM2.5, PM10, NO_2, CO, O_3	6 months

The Table.7 illustrates the diversity and volume of the dataset, combining both empirical and simulated sensor data to evaluate the robustness and generalizability of the proposed framework.

For comparative evaluation, three methods: LSTM-Based Prediction, GRU-Based Forecasting and CNN-LSTM Hybrid Models.

Table.8. Comparative Performance of Existing Methods and Proposed Transformer Model Across Sensor Rounds

No. of Sensors	Method	MAE (μg/m³)	RMSE (μg/m³)	R^2	MAPE (%)	Prediction Latency (ms)
10	LSTM	6.2	7.9	0.81	8.5	12
	GRU	5.9	7.6	0.83	8.1	10
	CNN-LSTM	5.1	6.5	0.87	7.0	18
	Proposed Transformer	4.3	5.2	0.91	5.8	15
20	LSTM	6.5	8.2	0.80	8.8	15
	GRU	6.1	7.8	0.82	8.3	12
	CNN-LSTM	5.4	6.8	0.86	7.3	20
	Proposed Transformer	4.5	5.4	0.90	6.0	17
30	LSTM	6.8	8.5	0.79	9.0	18
	GRU	6.3	8.0	0.82	8.5	14
	CNN-LSTM	5.6	7.0	0.85	7.5	23
	Proposed Transformer	4.7	5.6	0.89	6.2	18
40	LSTM	7.0	8.8	0.78	9.3	20
	GRU	6.5	8.2	0.81	8.8	16
	CNN-LSTM	5.8	7.2	0.84	7.8	25
	Proposed Transformer	4.9	5.8	0.88	6.5	20
50	LSTM	7.2	9.0	0.77	9.5	22
	GRU	6.7	8.4	0.80	9.0	18
	CNN-LSTM	6.0	7.5	0.83	8.0	28
	Proposed Transformer	5.1	6.0	0.87	6.8	22

Table.9. Comparative Performance of Existing Methods and Proposed Transformer Model on Real and Synthetic Datasets

Dataset	Method	Pollutant	MAE	RMSE	R^2	MAPE (%)	Prediction Latency (ms)
Real-World Sensor Data	LSTM	PM2.5	6.1	7.8	0.82	8.5	12
	GRU		5.8	7.5	0.84	8.1	10
	CNN-LSTM		5.0	6.4	0.87	7.0	18
	Proposed Transformer		4.2	5.1	0.91	5.8	15
	LSTM	PM10	7.0	8.5	0.81	9.0	12
	GRU		6.6	8.1	0.83	8.5	10
	CNN-LSTM		5.8	7.0	0.86	7.2	18
	Proposed Transformer		4.8	5.7	0.90	6.3	15
	LSTM	NO_2	4.5	5.8	0.80	7.5	12
	GRU		4.2	5.5	0.82	7.1	10
	CNN-LSTM		3.8	5.0	0.85	6.5	18
	Proposed Transformer		3.2	4.2	0.90	5.0	15
	LSTM	CO	0.9	1.2	0.79	6.5	12
	GRU		0.8	1.1	0.81	6.0	10
	CNN-LSTM		0.7	0.9	0.85	5.2	18
	Proposed Transformer		0.6	0.8	0.90	4.5	15
	LSTM	O_3	5.5	6.8	0.78	7.8	12
	GRU		5.2	6.4	0.80	7.3	10
	CNN-LSTM		4.7	5.5	0.84	6.5	18
	Proposed Transformer		3.9	4.6	0.89	5.5	15
Synthetic CRSN Simulation	LSTM	PM2.5	6.5	8.2	0.80	8.8	15
	GRU		6.1	7.8	0.82	8.3	12
	CNN-LSTM		5.4	6.8	0.86	7.3	20

	Proposed Transformer		4.5	5.4	0.90	6.0	17
	LSTM	PM10	7.2	8.7	0.79	9.2	15
	GRU		6.7	8.2	0.81	8.7	12
	CNN-LSTM		6.0	7.5	0.84	8.0	20
	Proposed Transformer		5.1	6.0	0.87	6.8	17
	LSTM	NO_2	4.8	6.0	0.78	7.9	15
	GRU		4.4	5.7	0.80	7.4	12
	CNN-LSTM		3.9	5.2	0.83	6.8	20
	Proposed Transformer		3.3	4.5	0.88	5.3	17
	LSTM	CO	1.0	1.3	0.77	6.7	15
	GRU		0.9	1.2	0.79	6.2	12
	CNN-LSTM		0.8	1.0	0.83	5.5	20
	Proposed Transformer		0.6	0.8	0.88	4.6	17
	LSTM	O_3	5.8	7.1	0.76	8.0	15
	GRU		5.4	6.6	0.78	7.6	12
	CNN-LSTM		4.9	5.7	0.82	6.8	20
	Proposed Transformer		4.0	4.8	0.87	5.6	17

The performance evaluation of the proposed transformer-based air pollution monitoring framework was conducted using both real-world sensor data and synthetic CRSN simulations. The experimental results, summarized in Table.3 and Table.4, demonstrate the superior predictive capability and robustness of the transformer model compared to conventional methods including LSTM, GRU, and CNN-LSTM. Across all experiments, the transformer model consistently achieved lower MAE and RMSE values, higher R^2, and lower MAPE, highlighting its effectiveness in modeling complex spatiotemporal dependencies among multiple pollutants.

Analyzing the sensor round experiments presented in Table.3, it is evident that as the number of sensors increases from 10 to 50, all models experience a gradual increase in MAE and RMSE due to the higher data complexity and potential sensor noise. For instance, the LSTM model's MAE increased from 6.2 μg/m³ to 7.2 μg/m³, while the GRU model's MAE increased from 5.9 μg/m³ to 6.7 μg/m³. In comparison, the transformer's MAE increased only from 4.3 μg/m³ to 5.1 μg/m³, maintaining superior predictive accuracy even under larger sensor deployments. Similarly, RMSE values demonstrate a similar trend, with the transformer consistently showing the lowest values across all sensor rounds (5.2–6.0 μg/m³), indicating better handling of outliers and extreme pollutant variations.

R^2 values also highlight the transformer's robustness, maintaining values above 0.87 even for 50 sensors, whereas LSTM and GRU drop below 0.80 in high-density scenarios. This suggests that the transformer effectively captures long-range temporal dependencies and inter-sensor spatial correlations, which conventional recurrent models struggle to handle. MAPE values, reflecting relative prediction errors, remained below 7% for the transformer across all sensor rounds, compared to up to 9.5% for LSTM at 50 sensors, indicating superior reliability in urban monitoring contexts. Additionally, the prediction latency for the transformer remained competitive (15–22 ms), slightly higher than GRU but substantially lower than CNN-LSTM, that shows its feasibility for real-time applications in IoT networks.

The results from the pollutant-specific evaluation across real-world and synthetic datasets, summarized in Table.4, further reinforce these conclusions. For PM2.5, a critical pollutant with high health impact, the transformer achieved MAE values of 4.2 μg/m³ (real) and 4.5 μg/m³ (synthetic), significantly outperforming LSTM (6.1–6.5 μg/m³) and GRU (5.8–6.1 μg/m³). PM10 results show a similar trend, with transformer MAE values of 4.8–5.1 μg/m³, reflecting its ability to handle particulate pollution, which exhibits both local and regional variability. For gaseous pollutants like NO_2, CO, and O_3, the transformer also demonstrates consistent superiority. For example, NO_2 MAE reduced to 3.2 μg/m³ in real-world datasets, compared to 4.5 μg/m³ for LSTM, while CO MAE reached only 0.6 ppm versus 0.9 ppm for LSTM. These reductions are significant in real-world monitoring, where small deviations can affect health risk assessment and policy decisions.

R^2 values for all pollutants consistently remained above 0.87 for the transformer, compared to 0.76–0.84 for conventional models, confirming enhanced model explainability and fit. MAPE improvements are also notable, with transformer errors reduced by approximately 20–30% relative to LSTM and GRU across all pollutants. The low prediction latency (15–17 ms) ensures that the system can generate timely air quality alerts for both high-density sensor networks and real-world urban monitoring, which is crucial for actionable decision-making.

Another key observation from the synthetic CRSN simulation is the model's ability to maintain performance under increased network density. As the number of synthetic sensors increased, conventional models exhibited greater degradation in MAE and RMSE due to difficulty in modeling the higher dimensional data. In contrast, the transformer maintained relatively stable performance, that shows its scalability and suitability for smart city implementations where high-density CRSN deployments are increasingly common. The self-attention mechanism allows the model to weigh contributions from all sensors dynamically, effectively mitigating noise and redundancy, which explains the improved R^2 and reduced MAPE even in dense networks.

The experimental analysis confirms that the transformer-based framework provides a robust and scalable solution for IoT-enabled air pollution monitoring. The model's ability to consistently achieve lower error rates, maintain high R^2 values, and provide real-time predictions across multiple pollutants and sensor densities highlights its practical applicability in urban environmental monitoring and public health management. The comparative analysis validates the significant advantages of integrating self-attention mechanisms over conventional recurrent architectures for both temporal and spatiotemporal pollution prediction.

5. CONCLUSION

In this study, a transformer-based deep learning framework was proposed for real-time air pollution monitoring using IoT-enabled Cognitive Radio-Based Sensor Networks. Experimental evaluation on both real-world sensor datasets and synthetic CRSN simulations demonstrated that the proposed model significantly outperforms conventional methods including LSTM, GRU, and CNN-LSTM. Across all sensor densities and pollutants, the transformer achieved lower MAE and RMSE, higher R^2, and reduced MAPE, while maintaining competitive prediction latency suitable for real-time applications. The results confirm that the self-attention mechanism effectively captures long-range temporal dependencies and inter-sensor spatial correlations, ensuring robustness and scalability in dense urban environments. By providing accurate and timely pollutant forecasts, the proposed framework facilitates informed decision-making for urban planning, public health advisories, and environmental management. This study establishes transformer-based modeling as a promising solution for next-generation IoT-enabled air quality monitoring systems that addresses the limitations of traditional recurrent approaches and advancing the state-of-the-art in smart environmental monitoring.

REFERENCES

[1] A.J. Al Ali and M.M.A. Alhaidery, "Machine Learning Techniques for Anomaly Detection in IoT and WSN: A Review", *Journal of Al-Qadisiyah for Computer Science and Mathematics*, Vol. 17, No. 2, pp. 229-240, 2025.

[2] Y. Zhao and D. Niyato, "TranDRL: A Transformer-Driven Deep Reinforcement Learning Enabled Prescriptive Maintenance Framework", *IEEE Internet of Things Journal*, Vol. 57, No. 1, pp. 1-14, 2024.

[3] W. El Gadal and S. Ganti, "Federated Secure Intelligent Intrusion Detection and Mitigation Framework for SD-IoT Networks using ViT-GraphSAGE and Automated Attack Reporting", *Proceedings of International Conference on New Technologies, Mobility and Security*, pp. 317-325, 2025.

[4] S.E. Bibri and J. Huang, "Generative AI of Things for Sustainable Smart Cities: Synergizing Cognitive Augmentation, Resource Efficiency, Network Traffic, Cybersecurity, and Anomaly Detection for Environmental Performance", *Sustainable Cities and Society*, Vol. 133, pp. 1-6, 2025.

[5] P. Sajjadi, F. Dinmohammadia and M. Shafiee, "Machine Learning in Prognostics and System Health Management of Cyber-Physical Systems: A Review", *IEEE Access*, Vol. 13, pp. 1-15, 2025.

[6] R. Aburukba and K. El-Fakih, "Optimizing Predictive Maintenance in Industrial IoT Cloud using Dragonfly Algorithm", *IEEE Internet of Things Journal*, Vol. 23, No. 1, pp. 1-17, 2025.

[7] R. Priyakanth, N.S. Krishna, P.R. Kumar and B. Deekshitha, "Visual Simulation of IoT-based Smart Cities inclining Towards the Attainment of Sustainable Development Goals", *Proceedings of International Conference on Modern Sustainable Systems*, pp. 978-982, 2025.

[8] H.K. Adli, M.A. Remli and M.S. Mohamad, "Recent Advancements and Challenges of a IoT Application in Smart Agriculture: A Review", *Sensors*, Vol .23, No. 7, pp. 3752-3759, 2023.

[9] A.M. Ashwini, N. Sridevi, N. Krishna, N.V. Babu and A. Kumar, "Enhancing Image Dehazing Efficiency with Dual Colour Space Attentional Deep Networks", *International Journal of Sensors, Wireless Communications and Control*, Vol. 23, No. 2, pp. 1-14, 2024.

[10] M. Zhou, D. Yang, X. Qiu and Y. Li, "Visual Geolocalization for Aerial Vehicles via Fusion of Satellite Remote Sensing Imagery and Its Relative Depth Information", *Remote Sensing*, Vol. 17, No. 13, pp. 2291-2306, 2025.

[11] W.A. Jabbar, W. Wu and M.A. De Oliveira, "LoRaWAN-based IoT System Implementation for Long-Range Outdoor Air Quality Monitoring", *Internet of Things*, Vol. 19, pp. 100540-100556, 2022.

[12] C.S. Ranganathan, G. Brindha, N. Mohankumar and S. Murugan, "Cloud-Combined Cognitive Radio-Based Sensor Networks for Dynamic Household Air Pollution Management", *Proceedings of Asia Pacific Conference on Innovation in Technology*, pp. 1-6, 2024.