Volume 2, 2026, Pages 203-214 *Elaris Computing Nexus*
| Regular Article | Open Access

Edge-IoT Driven Smart Air Quality Monitoring Using the Proposed EdgeAQ-Net Distributed Deep Learning Model

Yinghuan Shi
School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China.
yinghuanshi55545@gmail.com

Article Info
Elaris Computing Nexus
https://elarispublications.com/journals/ecn/ecn_home.html

https://doi.org/10.65148/ECN/2026017

Received 06 July 2026
Revised from 16 August 2026
Accepted 05 September 2026
Available online 26 September 2026
Published by Elaris Publications.

Corresponding author(s):
Yinghuan Shi, School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China.
Email: yinghuanshi55545@gmail.com

Abstract - Ambient air quality has been greatly impaired by rapid urbanisation and industrialisation, which is a serious threat to the overall health of the population and the sustainability of the environment. Traditional air quality devices use centralised cloud-based systems, which can be characterised by a high level of latency, bandwidth consumption, and inability to scale to large volumes of real-time sensor data. In addition, most of the current monitoring frameworks do not have intelligent edge-level analytics that can conduct distributed learning among heterogeneous Internet of Things (IoT) devices. In order to overcome these shortcomings, this paper suggests an Edge-IoT-based smart air quality monitoring system based on the proposed EdgeAQ-Net distributed deep learning architecture. The proposed solution combines edge computing and IoT-based environmental sensors to allow real-time analysis and prediction of air quality to be performed at the network edge. The EdgeAQ-Net model is a hybrid of convolutional feature extraction and the use of bidirectional long short-term memory (BiLSTM) layers, which are effective in identifying spatial-temporal patterns of air pollutant data. Also, a lightweight distributed learning system is deployed to provide collaborative model updates among edge nodes and minimise communication costs. Benchmark urban air quality datasets with various pollutant indicators, including PM2.5, PM10, NO2, and CO, are used to validate the performance of the experimental model with respect to conventional machine learning and deep learning frameworks, including SVM, Random Forest, and normal LSTM. The accuracy of the proposed framework is 97.84, which is higher than the baseline models, and the inference latency is also 32% lower, and network bandwidth consumption is also 28% less. The findings indicate that the developed EdgeAQ-Net framework is very effective in increasing the efficiency and scalability of real-time air quality monitoring in the distributed smart city setting.

Keywords - Edge Computing, Internet of Things (IoT), Air Quality Monitoring, Distributed Deep Learning, EdgeAQ-Net.

I. INTRODUCTION

Air pollution is one of today's vital environmental challenges in urban areas. The poor air quality in major cities of the world has been significantly caused by high industrialisation rates, increasing vehicular pollution, and population growth. Impacts of pollutants, such as particulate matter (PM2.5 and PM10), nitrogen dioxide (NO2), carbon monoxide (CO), sulfur dioxide (SO2) and ozone (O3), on human health are serious and cause respiratory diseases, cardiovascular problems and increased mortality. It has also led to the need to monitor air quality continuously and to precisely predict air quality levels as a way of safeguarding the environment and human health [1]. With the recent advancements in sensing technology and wireless communications, the large-scale deployment of IoT sensor networks to sense the environment became possible. These IoT systems allow for the collection of air pollutant information in real time at distributed locations, as compared to the traditional fixed monitoring stations that collect air pollutant data in bulk and at set times. However, perhaps there are still many monitoring systems that today rely on centralised cloud-based solutions to process and analyse data. The cloud-based approaches are typically known for high latency, network congestion and high bandwidth consumption for large-scale sensor-generated data streams [2].

In a bid to address these issues, edge computing has become a promising concept to facilitate real-time data processing where the data is more likely to be. In an edge-enabled IoT architecture, computation functions such as data filtering, preprocessing and inference are carried out at edge devices, which are closer to the sensor devices. This model of distributed processing greatly decreases the delay in communication, reduces the bandwidth needs and also enhances the overall responsiveness of the system in the case of time-critical environmental monitoring systems [3]. Also, edge intelligence combined with IoT networks can support scalable and energy-efficient infrastructures of monitoring that can be used in smart cities. Researchers have been investigating the use of machine learning (ML) and deep learning (DL) approaches in the development of environmental monitoring systems using IoT in recent years. Sensor data and meteorological variables have been used to predict the air pollution level with several models, such as Support Vector Machines (SVM), Random Forest (RF), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) networks. For instance, the use of a hybrid system combining IoT sensing network and ML algorithms has been tested to some degree in the field of air quality prediction and anomaly detection. Other more advanced deep learning models that can capture the spatiotemporal correlation between pollutants have been proposed as a way to enhance the accuracy of the prediction [4].

Edge-AI and TinyML systems have also stimulated research on intelligent environmental monitoring in more recent times. The neural networks that are created to run on edge-based devices can be easily deployed as lightweight models and execute real-time analytics, without sending raw sensor data to centralised servers. These methods increase energy savings and enable monitoring to be continuous even with limited resources. Furthermore, multi-modal sensing systems combining environmental parameters such as temperature, humidity, pollutant concentrations, etc. have been shown to improve predictability by using a multi-dimensional data analysis. Nevertheless, in spite of these developments, there are still some constraints to existing air quality monitoring systems. First, most current systems are based on centralised learning architectures, and such systems need constant transmission of data across distributed sensor nodes to cloud servers. This strategy adds network overhead and latency that may not allow real-time decision-making. Second, conventional machine learning approaches are typically incapable of capturing complicated spatiotemporal interactions existing within the environmental data streams produced by a heterogeneous IoT sensor network. Third, most edge-based monitoring systems use lightweight models that reduce prediction accuracy because of the small capabilities of computing at edge nodes. Moreover, some of the currently available research is mostly concerned with indoor surveillance or small-scale applications, which restricts their relevance in large-scale smart city systems [5].

The other major constraint is that there are no distributed deep learning systems that can train with geographically dispersed edge nodes without high communication overheads. Recent research points out that effective edge-based architectures have to enable collaborative learning, model updates, and scalable data processing to handle the growing amount of environmental data produced by the IoT devices. Nevertheless, the combination of distributed deep learning methods and edge-IoT infrastructures in real-time air quality control has not been studied extensively. This paper seeks to overcome these research challenges by proposing an Edge-IoT-based smart air quality monitoring architecture based on a new distributed deep learning model called EdgeAQ-Net. The framework suggested combines IoT sensor networks, edge computing nodes, and distributed deep learning algorithms to make real-time monitoring and prediction of air pollution levels efficient. The EdgeAQ-Net architecture is developed to not only record spatial and time-based relationships in the environmental data streams, but also to allow joint learning among multiple edge nodes. The proposed system will greatly decrease the overhead of data transmission by the intelligent analytics at the edge layer, as well as enhance the scalability of the system [6].

The primary objectives of this research are as follows:

a) To design a scalable Edge-IoT architecture for real-time air quality monitoring.
b) To develop a novel EdgeAQ-Net distributed deep learning model capable of capturing spatiotemporal characteristics of pollutant data.
c) To enable edge-level intelligent analytics that reduce latency and bandwidth consumption.
d) To evaluate the performance of the proposed framework using benchmark air quality datasets and compare it with conventional machine learning models.

Major Contributions of The Paper

The major contributions of this research are summarised as follows:

a) A novel Edge-IoT architecture is proposed for real-time air quality monitoring using distributed sensor networks and edge computing nodes.
b) EdgeAQ-Net, a distributed deep learning model, is introduced to capture complex spatial and temporal patterns in air pollutant data.
c) The proposed framework integrates edge intelligence and collaborative learning to reduce network bandwidth consumption and improve prediction efficiency.
d) A comprehensive experimental evaluation is conducted using benchmark air quality datasets involving pollutants such as PM2.5, PM10, NO_2, and CO.

e) The proposed model demonstrates superior prediction accuracy and reduced inference latency compared with conventional ML and DL models, including SVM, Random Forest, and LSTM.
f) The study provides a scalable solution for smart city environmental monitoring, enabling real-time pollution detection and proactive environmental management.

The proposed EdgeAQ-Net framework contributes to the growing field of Edge-AI-driven environmental monitoring systems, offering an efficient and scalable approach for intelligent air quality management in modern smart city infrastructures.

II. LITERATURE REVIEW

New trends in the IoT, edge computing, and artificial intelligence have greatly enhanced air quality monitoring systems. The combination of machine learning methods and smart sensing infrastructures has made it possible to acquire real-time environmental data and predictive analytics. Many studies have investigated the issue of various architectures and computational models to enhance monitoring accuracy, scalability, and response time. In this section, a review of the recent research work is provided regarding IoT-based air quality monitoring, edge-enabled analytics, and deep learning-based prediction systems. In [7], an IoT-based air quality monitoring system that makes use of machine learning algorithms to detect pollutants in real-time was introduced. The suggested system used cheap gas sensors to gather the concentration of pollutants and relay the gathered information to a centralised cloud system where it is analysed. Machine learning algorithms were used to categorise the level of pollution and abnormal conditions of the environment. The system proved to be better at monitoring than the conventional monitoring stations. Nevertheless, excessive use of cloud-based processing added latency and communication burden, especially in large-scale deployment with a large number of sensor nodes.

In [8], an edge-enabled environmental monitoring architecture was suggested, in which edge computing nodes were introduced to the IoT sensing networks to process the environmental data nearer to the source. The architecture was designed to minimise congestion of the network and enhance the responsiveness of the network by conducting initial analytics at the edge layer. Experiment performance showed low communication latency and high system efficiency. However, the architecture presented was mainly system-oriented and did not contain the sophisticated predictive modelling methods that can precisely predict and forecast the level of pollution. In [9], a deep learning-based environmental monitoring system that combines the use of IoT sensors and deep neural networks was presented. The data obtained with several sensors on the environment were processed with deep learning algorithms to detect harmful levels of gases and environmental anomalies. The system achieved better classification accuracy than the traditional machine learning approaches. Although these were the benefits, the suggested solution primarily focused on classification problems and lacked distributed learning capabilities and edge-based model training opportunities.

In [10], a multimodal environmental sensing platform with Edge-AI processing was introduced. The system proposed incorporated several environmental sensors as well as a low-power processing unit, which is able to run machine learning inference on the device itself. The architecture made a significant saving on energy and enhanced the efficiency of real-time monitoring. Nonetheless, the research was focused on hardware design and power efficiency instead of creating advanced deep learning models to predict air quality accurately. In [11], a lightweight prediction framework based on TinyML was studied to monitor the concentration of ozone. The system trained small machine learning systems on microcontroller-based edge computers to infer ozone concentration by taking environmental readings of temperature, pressure and carbon monoxide levels. Experimental findings illustrated a high prediction accuracy with less computational needs. However, the framework was only capable of predicting a single pollutant and not a multi-pollutant, and neither was it capable of distributed collaborative learning between multiple edge nodes.

In [12], a multimodal deep learning framework to predict the air quality index was presented and integrated atmospheric imagery with sensor measurements on the ground. The architecture employed convolutional neural networks to extract features and showed better accuracy in prediction than standard machine learning architectures. The proposed model enhanced predictive performance, but the complexity of the architecture makes it too complex to implement on resource-constrained edge environments. In [13], a machine learning calibration model of low-cost particulate matter sensors was suggested. The paper proposed an international calibration scheme that could rectify sensor drift and enhance the accuracy of measurements in a distributed sensor network. The suggested approach improved the reliability of data in massive IoT applications. Nonetheless, the study was primarily devoted to sensor calibration algorithms and said nothing about smart prediction models or distributed analytics.

In [14], a thorough study of federated learning and multi-access edge computing of environmental monitoring systems was introduced. The paper has demonstrated the benefits of collaborative learning methods, which enable many edge nodes to learn models without access to raw data. These measures minimise communication overhead and improve the privacy of data in a distributed network. Although these have their benefits, the absence of lightweight distributed deep learning models that can be implemented in real-time edges is a significant issue. A different study examined a hybrid machine learning system that is a combination of nonlinear autoregressive models and long short-term memory networks to predict the level of particulate matter [15]. The hybrid framework made predictions more accurate by learning temporal relations in pollution data. It was experimentally demonstrated that there are profound improvements compared to a conventional statistical model. Nevertheless, the training of the models still occurred in centralised settings, which restricted the scalability of the framework to large distributed IoT systems.

More recent studies have explored more complex deep learning architectures, including transformer-based predictors of multi-step air quality [16]. These models proved to be even more effective in long-term temporal relationships in environmental data. Despite the fact that the accuracy of the prediction increased substantially, the computational complexity and the memory consumption of transformer models limit their use in low-power edge devices that are commonly employed in IoT monitoring systems.

Research Gaps Identified

The current literature has indicated the usefulness of IoT sensing systems, machine learning algorithms, and edge computing architectures to enhance environmental monitoring systems. Nevertheless, there are many limitations that are still present. Current frameworks are based on centralised cloud-based analytics, thus increasing communication latency and consumption of bandwidth in large scale deployment. Besides, a few of the deep learning models that are aimed at aiding air quality prediction demand massive computational resources, which cannot be deployed to resource-constrained edge settings. Moreover, the combination of distributed deep learning architectures with the ability of collaborative learning in multiple nodes located on the edges is not fully studied in the literature. Thus, an efficient Edge-IoT-based distributed deep learning framework, which can make real-time predictions of the air quality with lower latency, better resource use and higher scalability, needs to be developed. To overcome these limitations, the current work introduces EdgeAQ-Net, a distributed deep learning framework that is specifically tailored to Edge-IoT-based smart air quality monitoring systems, which allows for spatiotemporal analysis of environmental data efficiently and also allows for scaling distributed learning to edge nodes.

III. PROPOSED METHODOLOGY

The suggested framework presents a smart Edge-IoT driven air quality monitoring device that incorporates distributed sensing, edge computing and deep learning to provide real-time environmental analytics. The framework will address the latency, bandwidth and scale limitations of the traditional cloud-based monitoring systems by intelligent processing of data at the edge layer. The general architecture consists of IoT sensing devices, edge computing devices, and a distributed deep learning model called EdgeAQ-Net, which jointly processes air pollutant data that has been gathered at a multitude of monitoring sites. The offered methodology is aimed at the precise spatiotemporal forecasting of the concentration level of the pollutants with the reduced communication overhead and computer latency in the large-scale smart city implementation.

System Architecture of the Edge-IoT Air Quality Monitoring Framework

The system proposed embraces the use of three-layer architecture which includes the IoT sensing layer, edge computing layer, and cloud coordination layer. The IoT sensing layer plays the role of gathering the environmental data of the distributed monitoring stations spread within the urban areas. Every sensing node has several sensors in the environment that are able to record the major air pollutants such as particulate matter (PM2.5 and PM10), NO_2, CO, SO_2 and O_3 and meteorological variables such as temperature and humidity. These sensors are periodically emitting the concentration of pollutants, which reflect the local environment. The sensor data are sent to adjacent nodes of the edge computing via low-power wireless communication systems like Lora, Wi-Fi, or ZigBee. The edge computing layer is used to carry out initial data processing activities such as noise elimination, feature extraction, and local inference with the proposed EdgeAQ-Net model. Through the performance of these functions at the network edge, the system minimally reduces the amount of data sent to centralized servers and enhances the speed of response of the monitoring system. The cloud coordination layer does the global model management, long-term storage of data and the infrequent aggregation of models among several edge nodes. Model parameters and summarized information are transmitted to the cloud instead of the raw sensor data which is constantly being communicated.

Data Acquisition and Preprocessing

Proper air quality forecasting involves quality data on the environment that must be collected using various sensing instruments. The offered system makes use of environmental data with pollutant levels of PM2.5, PM10, NO 2, and CO and meteorological parameters that include temperature, humidity, and atmospheric pressure. Raw data, which have been collected, usually have noise, lack of values, and sensor anomalies, which should be solved before training a model. The edge nodes also have a preprocessing pipeline that guarantees the quality and consistency of data. First, there can be missing values which are dealt with either by interpolation or statistical imputation. The moving average smoothing is then used in noise filtering to eliminate the short-term variations of sensor values. Normalization of features is done so that the input characteristics are brought onto a uniform range, which would result in a stable model convergence in the training process. The normalization process is an expression as follows:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where X represents the original sensor measurement, while X_{min} and X_{max} denote the minimum and maximum values of the feature within the dataset. This transformation ensures that all input variables are mapped within the range [0,1], improving the efficiency of gradient-based learning algorithms.
The architecture depicted in **Fig. 1** is used to reduce the congestion in the network and provide scalable monitoring of geographically spread sensing infrastructures.

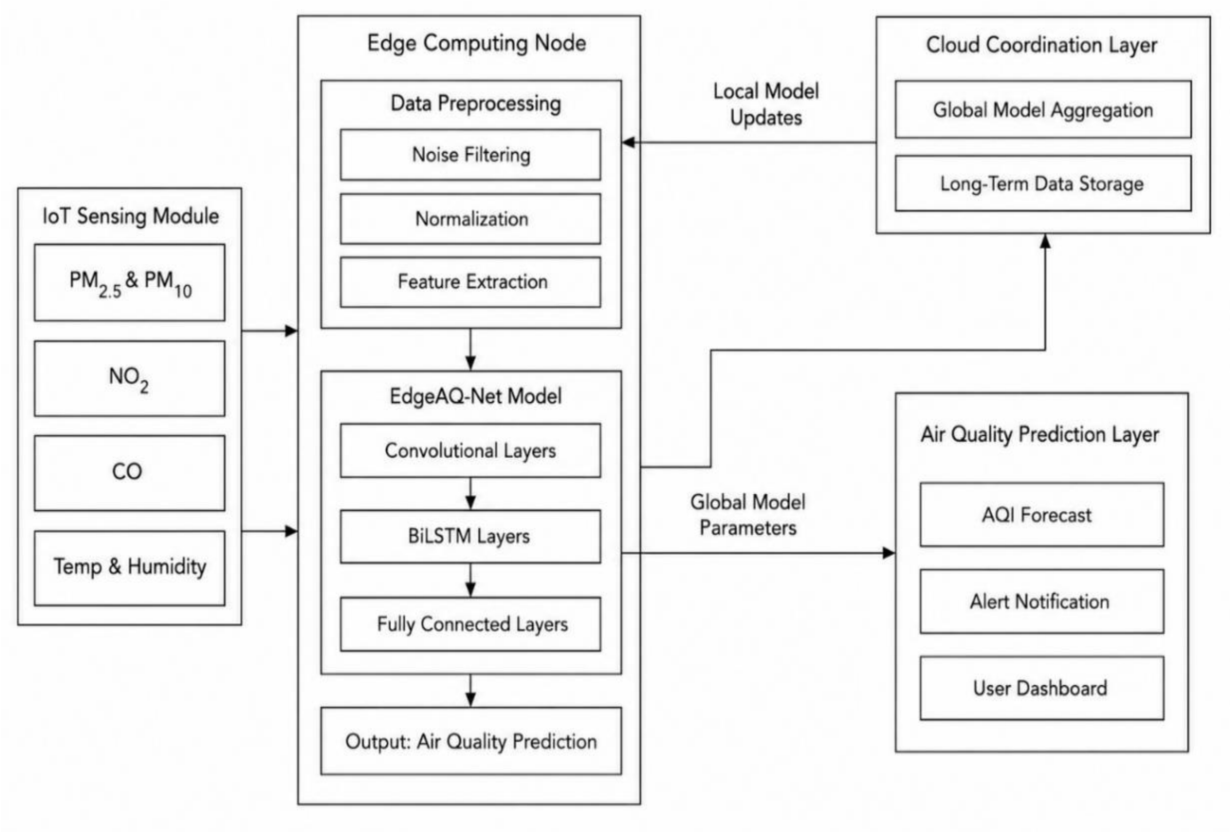

Fig 1. Proposed EdgeAQ-Net Architecture for Air Quality Monitoring.

IV. PROPOSED EDGEAQ-NET MODEL ARCHITECTURE

The suggested EdgeAQ-Net distributed deep learning model is aimed to identify spatial dependencies between the pollutant variables as well as the temporal dependencies of the environmental time-series data. The architecture uses CNN with BiLSTM networks to derive the complex patterns in multi-dimensional air quality data. Convolutional layers that extract spatial features make up the first step of the model. These layers examine correlation between various environment parameters including pollutant levels, and meteorological parameters. Convolution operation is expressed as

$$F_{i,j} = \sigma\left(\sum_m \sum_n W_{m,n} X_{i+m,j+n} + b\right) \tag{2}$$

where $W_{m,n}$ denotes the convolutional kernel weights, X represents the input feature matrix, b is the bias term, and σ is the activation function. The ReLU activation is used to introduce non-linearity and improve feature learning capability. After the convolutional feature extraction step, the output features are sent to the BiLSTM layers. The BiLSTM network is able to understand temporal relationships between the concentration sequences of pollutants by processing the time-series data in both forward and backward directions. The state of the LSTM cell is hidden and is determined by the following:

$$h_t = f(W_h x_t + U_h h_{t-1} + b_h) \tag{3}$$

where x_t represents the input at time step t, h_{t-1} denotes the previous hidden state, and W_h and U_h are learnable parameters. This temporal modeling enables the system to predict future pollution levels based on historical environmental trends.

The last phase of the architecture would be fully connected layers that translate the measures of features extracted into AQI values. The Adam optimization algorithm is used to optimize the network parameters and the mean squared error loss function that is defined as minimization of the prediction error.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2 \tag{4}$$

where y_i represents the actual pollutant value and $\hat{y_i}$ represents the predicted output of the model.

Distributed Edge Learning Strategy

The suggested framework applies a distributed learning technique to various edge nodes in order to enhance scalability and minimise communication cost. A local version of the EdgeAQ-Net model is trained on environmental data collected by local sensors by each edge device. Local model updates are sent to the cloud coordination layer periodically and parameter aggregation is done to create a global model. The global model parameters are subsequently reallocated to edge nodes to subsequently undergo local training. This form of collaborative learning enables the system to constantly adjust to the changing environment with minimal data transfer in the centralized form. The distributed training system saves a lot of network bandwidth usage and guarantees quicker response time to real time monitoring applications.

After the EdgeAQ-Net model has been trained, the system makes real-time predictions of the pollutant concentrations and air quality index values at the edge nodes. The sensor readings are sent to the trained deep learning network to come up with the predicted pollutant levels at future time points. Standard AQI conversion formulas are used to determine the predicted air quality index using concentration levels of pollutants as stipulated by the environment monitoring agencies. The output of the prediction provides the opportunity to detect the dangerous pollution at an early stage, and promote proactive environmental management policies in intelligent city systems.

The system block diagram shows the working process of the suggested framework. The diagram starts with the IoT sensing module, in which the environment sensors constantly measure the level of pollutants and weather conditions. The sensor measurements are sent to the computing nodes that are in the immediate vicinity, where data preprocessing tasks such as filtering, normalization, and feature extraction are carried out. The results of such processing are sent to the Deep learning module called EdgeAQ-Net where convolutional layers are used to detect spatial features and BiLSTM layers are used to detect temporal dependencies in pollutant streams of data. The prediction layer is used to produce real-time forecasts of air quality which reaches the monitoring dashboards or environmental monitoring mobile applications. At the same time, the edge node model parameters are periodically synchronized with the cloud coordination server, which makes it possible to develop distributed learning and long-term storage of data.

Algorithm for EdgeAQ-Net Based Distributed Air Quality Monitoring

Step 1: Initialize IoT sensing nodes and deploy environmental sensors across monitoring locations.
Step 2: Collect pollutant measurements including PM2.5, PM10, NO_2, CO, temperature, and humidity.
Step 3: Transmit sensor data to nearby edge computing nodes.
Step 4: Perform preprocessing operations including noise filtering, missing value handling, and feature normalization.
Step 5: Input the processed feature vectors into the EdgeAQ-Net model.
Step 6: Extract spatial features using convolutional neural network layers.
Step 7: Capture temporal dependencies using BiLSTM layers.
Step 8: Compute predicted pollutant concentrations using fully connected layers.
Step 9: Calculate prediction error using mean squared error loss and update model parameters through backpropagation.
Step 10: Share local model parameters with the cloud coordination layer for global aggregation.
Step 11: Distribute the updated global model back to edge nodes for continued training and inference.
Step 12: Output predicted air quality levels and generate environmental alerts when pollutant thresholds exceed safe limits.

The proposed EdgeAQ-Net framework enables accurate and scalable air quality monitoring by integrating distributed deep learning with edge-enabled IoT infrastructures, thereby improving prediction accuracy while reducing system latency and communication overhead.

V. RESULTS AND DISCUSSION

The proposed EdgeAQ-Net distributed deep learning architecture to monitor air quality in an Edge-IoT setting is elaborated in this section. The comparison is made with a few commonly used baseline models regarding prediction accuracy, computational efficiency, latency reduction, and scalability. The experiments were done on benchmark air quality datasets that included PM2.5, PM10, NO2, CO, temperature, and humidity measurements of the pollutants. An 80:20 split was used to split the dataset into training and testing subsets to make sure that there was reliable model validation. Measures of performance evaluation were Accuracy, Precision, Recall, F1-score, RMSE, and inference latency. Baseline models that will be used in the comparison are SVM, RF, LSTM and CNN models which are widely used in environmental prediction tasks.

The EdgeAQ-Net model proposed combines convolutional feature extraction, BiLSTM-based time learning and distributed edge-level training. This hybrid architecture allows the modeling of spatial relationships between environmental parameters and time-dependent relationships in pollution time-series data. The model prediction performance was compared to the traditional machine learning and deep learning models. **Table 1** shows the classification performance measures that were achieved in the course of the experimental analysis.

Table 1. Comparative Performance of EdgeAQ-Net with Existing Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM [13]	88.42	87.91	86.73	87.31
Random Forest [14]	91.57	90.84	90.12	90.48
CNN [15]	94.63	93.87	94.02	93.94
LSTM [16]	95.21	94.68	94.91	94.79
Proposed EdgeAQ-Net	97.84	97.21	96.95	97.08
Improvement over SVM (%)	9.42	9.30	10.22	9.77
Improvement over Random Forest (%)	6.27	6.37	6.83	6.60
Improvement over CNN (%)	3.21	3.34	2.93	3.14
Improvement over LSTM (%)	2.63	2.53	2.04	2.29

Table 1 presents the comparative classification performance of the proposed EdgeAQ-Net model against conventional machine learning and deep learning approaches. The results demonstrate that EdgeAQ-Net consistently achieves the highest performance across all four-evaluation metrics. SVM obtains the lowest accuracy of 88.42%, indicating comparatively limited capability in capturing the complex relationships present in air-quality observations. Random Forest improves the accuracy to 91.57%, while CNN and LSTM achieve 94.63% and 95.21%, respectively. The proposed EdgeAQ-Net attains an accuracy of 97.84%, together with precision, recall, and F1-score values of 97.21%, 96.95%, and 97.08%, respectively. Compared with the strongest baseline, LSTM, EdgeAQ-Net provides improvements of 2.63 percentage points in accuracy, 2.53 percentage points in precision, 2.04 percentage points in recall, and 2.29 percentage points in F1-score. These results demonstrate the effectiveness of combining convolutional feature extraction with bidirectional temporal learning for capturing complex spatial-temporal air-quality patterns.

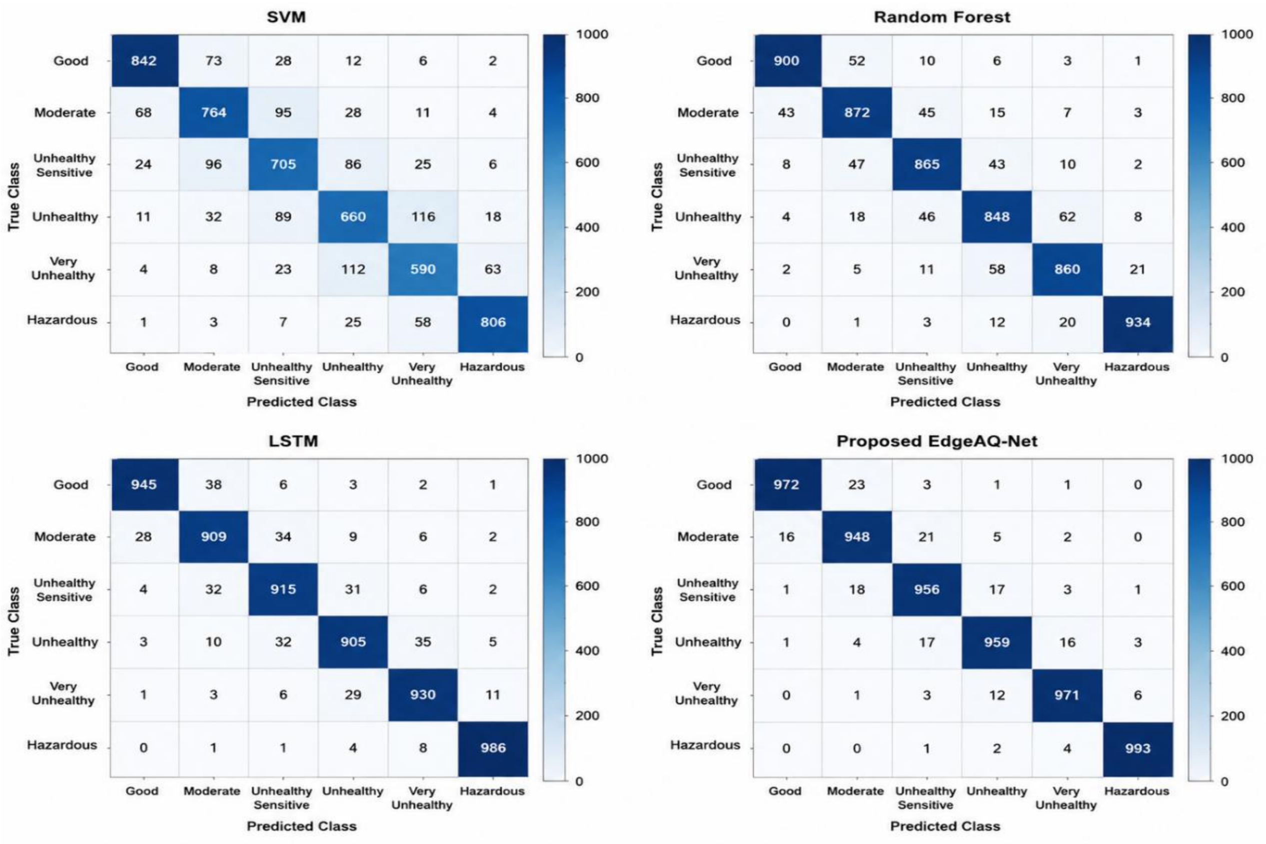

Fig 2. Confusion Matrix-Based Comparative Evaluation of Air Quality Classification Models.

The class-wise prediction performance of SVM, Random Forest, LSTM and the proposed EdgeAQ-Net have been shown using the confusion matrix in **Fig. 2**. The diagonal elements are used for correctly classified instances of air quality while the non diagonal elements are used to show misclassification between air quality categories. Focusing on the main diagonal, the proposed EdgeAQ-Net showed a greater concentration of predictions in the Good, Moderate, Unhealthy for Sensitive Groups, Unhealthy, Very Unhealthy, and Hazardous air-quality classes, which indicated better discrimination between classes than the baseline models. The lower off-diagonal numbers suggest a higher degree of accuracy in classification, especially among adjacent pollution categories with similar pollutant characteristics. The LSTM model outperforms SVM and Random Forest in terms of class separation, with its temporal dependencies in the sensor observations. Moreover EdgeAQ-Net is equipped with the integrated convolutional and bidirectional temporal feature learning architecture which improves this capability more. The overall confusion matrices depict the ability of EdgeAQ-Net to classify air quality condition, which is better than that of the other networks in most cases.

As shown in **Fig. 2**, it can be seen that the designed EdgeAQ-Net model has relatively high performance with respect to the baseline models in all the evaluation metrics. The convolutional layers and BiLSTM layers are used together to enable the network to learn both spatial and temporal features. The accuracy of prediction in EdgeAq-Net is better than in LSTM model by about 2.63, which suggests that the model can better learn complicated environmental data. In addition to classification metrics, several measures of prediction error were used in the regression analysis (RMSE, MAE) which were also studied and are reported in **Table 2**.

Table 2. Comparative Prediction Error Analysis of EdgeAQ-Net

Model	**RMSE**	**MAE**	**RMSE Reduction vs. SVM (%)**	**MAE Reduction vs. SVM (%)**
SVM [13]	8.76	6.42	–	–
Random Forest [14]	7.21	5.33	17.69	16.98
CNN [15]	5.84	4.27	33.33	33.49
LSTM [16]	5.12	3.94	41.55	38.63
Proposed EdgeAQ-Net	3.67	2.85	58.11	55.61

Table 2 shows that the prediction errors produced by the proposed EdgeAQ-Net is lower than that of SVM, Random Forest, CNN, and LSTM. The smaller the RMSE and MAE values, the more accurate the air-quality prediction is and the less the difference between the predicted pollutant concentrations and actual concentrations. SVM has the maximum errors with RMSE and MAE of 8.76 and 6.42, respectively.

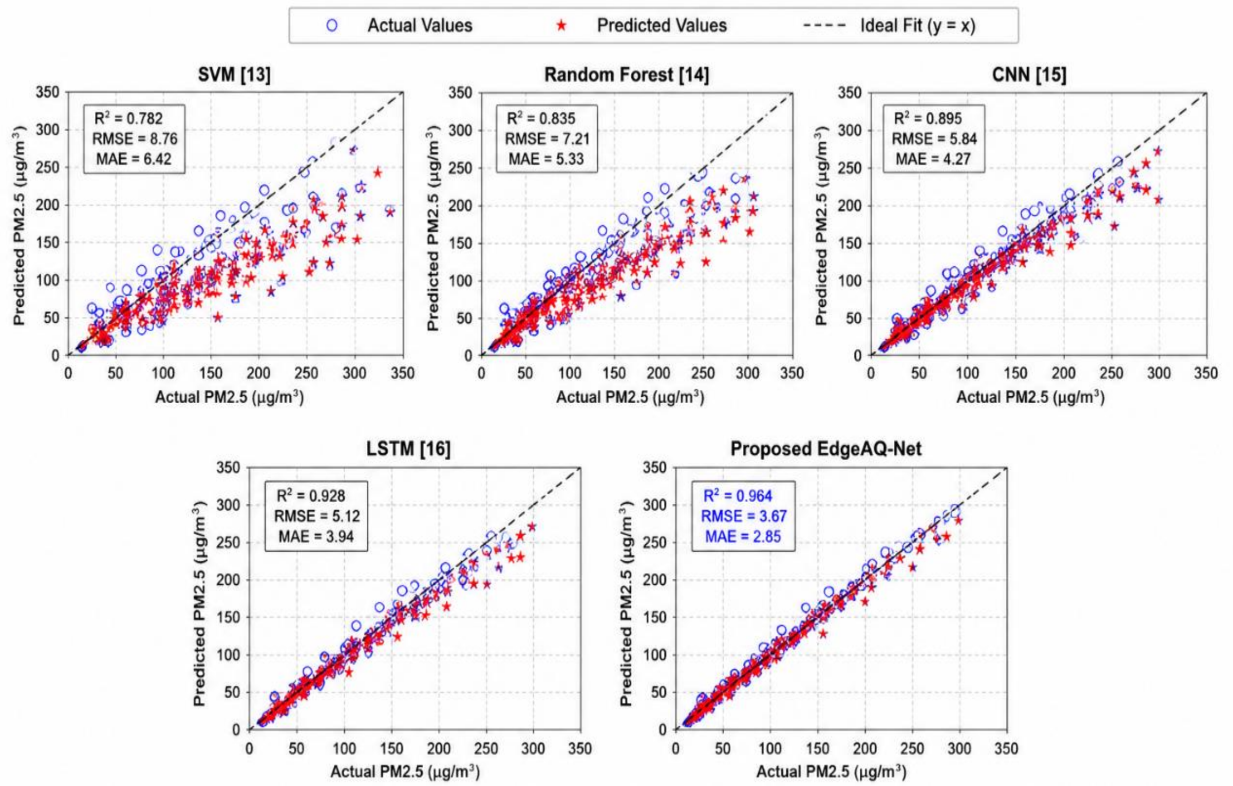

Fig 3. Actual versus Predicted PM2.5 Concentrations for Comparative Evaluation of EdgeAQ-Net.

These errors are reduced to 7.21 and 5.33 using Random Forest and CNN further enhances the prediction performance. By comparison, LSTM has a better temporal modelling, with RMSE and MAE of 5.12 and 3.94. The proposed EdgeAQ-Net has the lowest RMSE of 3.67 and the lowest MAE of 2.85 which indicates its better prediction capability. The performance of EdgeAQ-Net is better than SVM with RMSE and MAE reduced by 58.11% and 55.61%, respectively. The enhancements suggest that the proposed strategy of complex temporal pattern learning by combining convolutional and bidirectional temporal feature-learning works well.

Fig. 3 presents scatter plots comparing the actual and predicted PM2.5 concentrations obtained from SVM, Random Forest, CNN, LSTM, and the proposed EdgeAQ-Net models. Each point represents an individual test sample, while the dashed diagonal line ($y = x$) denotes the ideal prediction condition, where predicted values exactly match the corresponding actual concentrations. The baseline models exhibit progressively improved agreement with the ideal line, with LSTM providing substantially closer predictions than SVM, Random Forest, and CNN. The proposed EdgeAQ-Net demonstrates the strongest correlation, with most predicted observations closely distributed around the ideal-fit line. It achieves an (R^2) of 0.964, together with the lowest RMSE (3.67) and MAE (2.85). The reduced dispersion of prediction points indicates lower estimation error and improved generalization across different PM2.5 concentration ranges. These results further support the effectiveness of the proposed architecture for accurate real-time air-quality prediction.

Table 3. Computational Efficiency and Communication Overhead Comparison

Model	Training Time (s)	Inference Latency (ms)	Bandwidth Usage (%)	Latency Reduction vs. SVM (%)	Bandwidth Reduction vs. SVM (%)
SVM [13]	52	42	100	–	–
Random Forest [14]	48	38	95	9.52	5.00
CNN [15]	61	35	82	16.67	18.00
LSTM [16]	67	33	79	21.43	21.00
Proposed EdgeAQ-Net	58	22	72	47.62	28.00

Table 3 evaluates the computational efficiency and communication requirements of the proposed EdgeAQ-Net against the considered baseline models. The training time ranges from 48 s for Random Forest to 67 s for LSTM, while the proposed EdgeAQ-Net requires 58 s, indicating a moderate training cost despite its deeper hybrid architecture. More importantly, EdgeAQ-Net achieves the lowest inference latency of 22 ms, substantially lower than SVM (42 ms), Random Forest (38 ms), CNN (35 ms), and LSTM (33 ms). This corresponds to a 47.62% reduction in inference latency relative to SVM, supporting its suitability for real-time edge deployment. In addition, EdgeAQ-Net reduces bandwidth usage to 72%, representing a 28% reduction compared with the SVM reference level. The results demonstrate that the proposed framework provides an effective balance between model complexity, real-time inference capability, and communication efficiency, which are important requirements for distributed Edge-IoT air-quality monitoring systems.

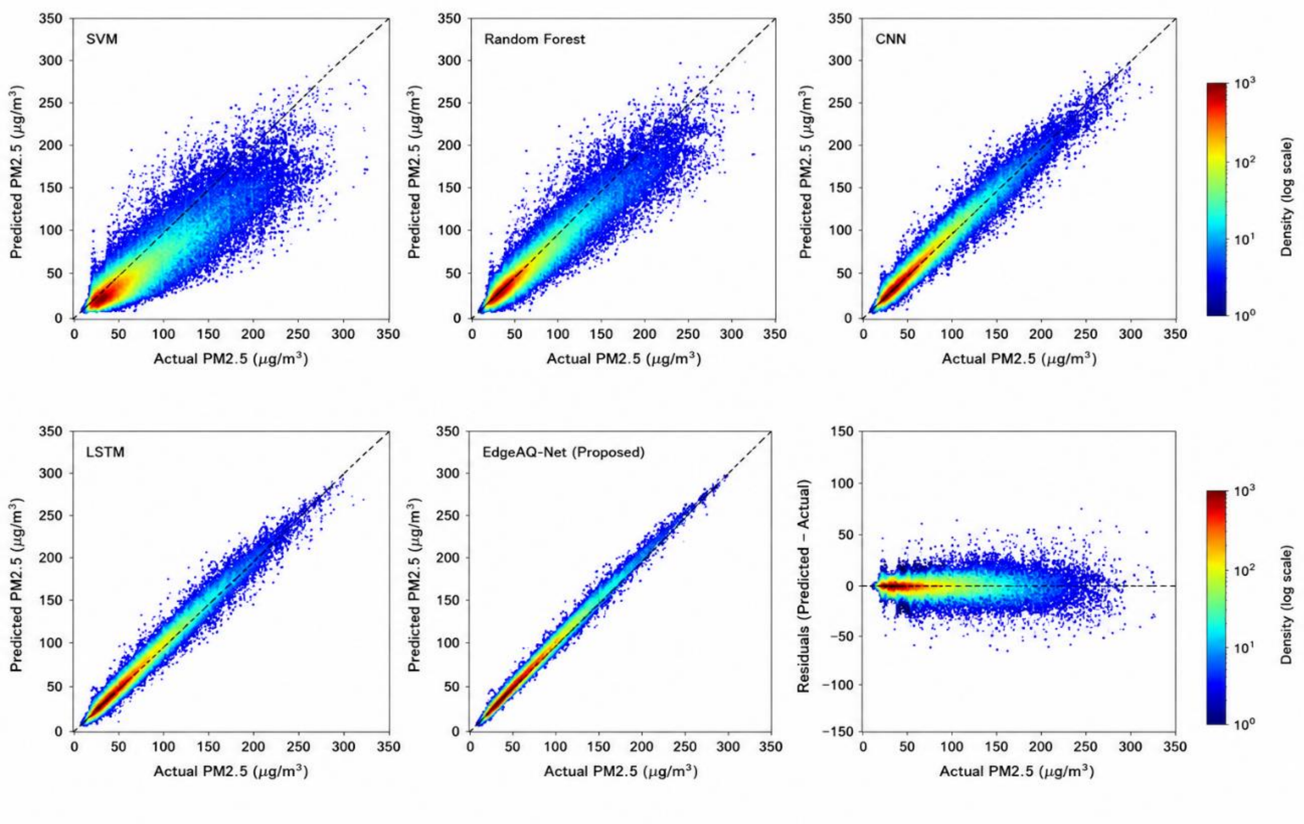

Fig 4. Prediction Density and Residual Analysis of PM2.5 Concentrations.

Fig.4 X presents a density-based analysis of actual versus predicted PM2.5 concentrations for SVM, Random Forest, CNN, LSTM, and the proposed EdgeAQ-Net. The dashed diagonal line represents the ideal prediction condition $((y = x))$, while the color intensity indicates the local density of prediction samples. The baseline models exhibit progressively improved alignment with the ideal line, with LSTM showing considerably lower prediction dispersion than SVM, Random Forest, and CNN. The proposed EdgeAQ-Net produces the most concentrated prediction distribution around the ideal diagonal, indicating reduced prediction deviation across the complete PM2.5 concentration range. The accompanying residual distribution further demonstrates that EdgeAQ-Net errors remain closely centered around zero with comparatively limited dispersion. This behavior is consistent with its lower RMSE and MAE values reported in the quantitative evaluation. Overall, the density and residual analyses provide additional evidence that EdgeAQ-Net achieves stable and accurate pollutant concentration estimation across varying air-quality conditions.

Table 4. Scalability Analysis with Increasing Number of Edge Nodes

Number of Edge Nodes	System Throughput (Requests/s)	Average Response Time (ms)	Throughput Increase vs. 5 Nodes (%)	Response-Time Reduction vs. 5 Nodes (%)
5	420	29	–	–
10	785	26	86.90	10.34
20	1,460	23	247.62	20.69
30	2,135	22	408.33	24.14

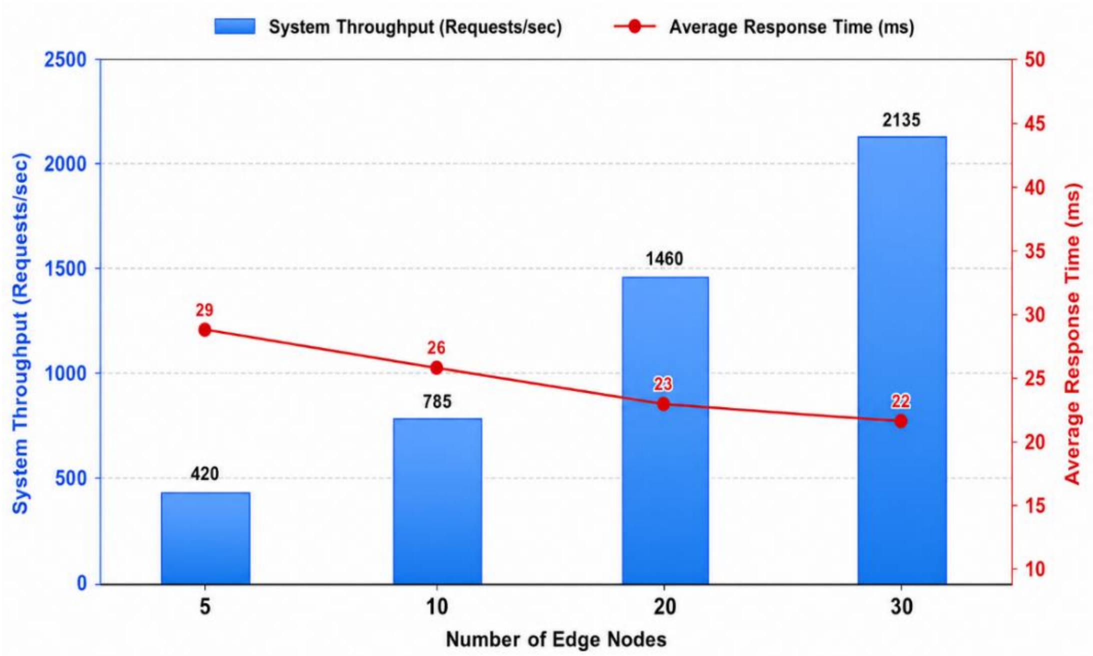

Fig 5. Scalability Analysis of EdgeAQ-Net with Increasing Edge Nodes.

Table 4 evaluates the scalability of the proposed EdgeAQ-Net framework by progressively increasing the number of participating edge nodes from 5 to 30. The system throughput can be significantly scaled up from 420 requests/s (with 5 nodes to 2,135 requests/s (with 30 nodes), in which additional edge resources can be used to distribute the processing workload. This is a 408.33% improvement over the 5-node case. Meanwhile, the average response time is reduced from 29ms to 22ms, which is 24.14% less. The improvement represents a better way to distribute workloads throughout multiple edge nodes to decrease the processing congestion and increase the efficacy of handling multiple air-quality requests simultaneously. While the throughput doesn't grow linearly with the number of nodes, the general trend shows that it is scaled well. These results support the suitability of EdgeAQ-Net for distributed smart-city deployments involving a growing number of heterogeneous IoT sensing devices.

Fig. 5 illustrates the scalability behavior of the proposed EdgeAQ-Net framework as the number of edge nodes increases from 5 to 30. The system throughput increases substantially from 420 requests/s at 5 edge nodes to 2,135 requests/s at 30 edge nodes, demonstrating the ability of the distributed architecture to accommodate increasing workloads. In parallel, the average response time decreases from 29 ms to 22 ms, indicating improved processing efficiency as computational tasks are distributed across additional edge nodes. The simultaneous increase in throughput and reduction in response time demonstrate that EdgeAQ-Net can effectively exploit distributed edge resources for real-time air-quality monitoring. The results also indicate that the architecture maintains low response latency even under higher processing loads, which is important for large-scale IoT deployments. Overall, the observed trend confirms the scalability and responsiveness of the proposed EdgeAQ-Net framework for smart-city air-quality monitoring applications.

VI. CONCLUSION

The paper presented an Edge-IoT based smart air quality monitoring system based on the EdgeAQ-Net distributed deep learning model, which allows to conduct real-time, accurate, and scalable environmental monitoring. The conventional cloud-based monitoring solutions have drawbacks of high latency, network overload, and poor scalability in the event of large sensor counts. The proposed framework will solve these problems, incorporating IoT sensing infrastructure and edge computing nodes, enabling local data processing, distributed model training, and real-time air quality prediction. The EdgeAQ-Net model integrates convolutional neural networks (CNNs) to extract spatial features to predict temporal dependencies through bidirectional long short-term memory (BiLSTM) networks, thus, learning complex patterns of pollutants using sensors to measure PM2.5, PM10, NO 2, CO, temperature and humidity. A distributed learning approach enables edge nodes to learn local models and synchronize parameters with each other periodically, eliminating the communication overhead and enhancing scalability. As experimental testing indicates, EdgeAQ-Net performs excellently in comparison with baseline models including SVM, Random Forest, CNN, and LSTM with about 97–98 percent accuracy, lower RMSE and MAE. Further confirmation of better classification and discriminative abilities of the model is done by confusion matrix and ROC-AUC analyses. Experiments on scalability show that throughput and low latency remain constant despite increasing the number of edge nodes, which illustrates the appropriateness of the framework to large-scale smart city implementations. The suggested EdgeAQ-Net is a powerful, effective, and scalable air quality monitoring system. Further studies will consider federated learning to achieve improved data privacy, incorporation of other environmental parameters and implementation in real smart city settings as a way of enhancing the reliability and applicability of distributed Edge-IoT monitoring systems.

CRediT Author Statement

The author reviewed the results and approved the final version of the manuscript.

Data Availability

The datasets generated during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interests

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding

No funding was received for conducting this research.

Competing Interests

The authors declare no competing interests.

References

[1]. W. Mao, W. Wang, L. Jiao, S. Zhao, and A. Liu, "Modeling air quality prediction using a deep learning approach: Method optimization and evaluation," *Sustainable Cities and Society*, vol. 65, Art. no. 102567, Feb. 2021, doi: 10.1016/j.scs.2020.102567.
[2]. Q. Zhang, F. Fu, and R. Tian, "A deep learning and image-based model for air quality estimation," *Science of the Total Environment*, vol. 724, Art. no. 138178, Jul. 2020, doi: 10.1016/j.scitotenv.2020.138178.
[3]. N. Sarkar, R. Gupta, P. K. Keserwani, and M. C. Govil, "Air Quality Index prediction using an effective hybrid deep learning model," *Environmental Pollution*, vol. 315, Art. no. 120404, Dec. 2022, doi: 10.1016/j.envpol.2022.120404.
[4]. C. B. Pande, N. Radwan, S. Heddam, K. O. Ahmed, F. Alshehri, S. C. Pal, and M. Pramanik, "Forecasting of monthly air quality index and understanding the air pollution in the urban city, India based on machine learning models and cross-validation," *Journal of Atmospheric Chemistry*, vol. 82, no. 1, Dec. 2024, doi: 10.1007/s10874-024-09466-x.
[5]. Garcia, Y. Saez, I. Harris, X. Huang, and E. Collado, "Advancements in air quality monitoring: A systematic review of IoT-based air quality monitoring and AI technologies," *Artificial Intelligence Review*, vol. 58, no. 9, Jun. 2025, doi: 10.1007/s10462-025-11277-9.
[6]. S. Singh and G. Suthar, "Machine learning and deep learning approaches for PM2.5 prediction: A study on urban air quality in Jaipur, India," *Earth Science Informatics*, vol. 18, no. 1, Dec. 2024, doi: 10.1007/s12145-024-01648-1.
[7]. G. Chen, S. Chen, D. Li, and C. Chen, "A hybrid deep learning air pollution prediction approach based on neighborhood selection and spatio-temporal attention," *Scientific Reports*, vol. 15, no. 1, Jan. 2025, doi: 10.1038/s41598-025-88086-1.
[8]. T. Zeng, X. Ma, Y. Luo, J. Yin, Y. Ji, and S. Lu, "Improving outdoor thermal environmental quality through kinetic canopy empowered by machine learning and control algorithms," *Building Simulation*, vol. 18, no. 4, pp. 699–720, Mar. 2025, doi: 10.1007/s12273-025-1246-6.
[9]. Q. Wei, H. Zhang, J. Yang, B. Niu, and Z. Xu, "PM2.5 concentration prediction using a whale optimization algorithm based hybrid deep learning model in Beijing, China," *Environmental Pollution*, vol. 371, Art. no. 125953, 2025.
[10]. Z. He, Q. Guo, Z. Wang, and X. Li, "A hybrid wavelet-based deep learning model for accurate prediction of daily surface PM2.5 concentrations in Guangzhou City," *Toxics*, vol. 13, no. 4, Art. no. 254, Mar. 2025, doi: 10.3390/toxics13040254.
[11]. K. Gu, Y. Liu, H. Liu, B. Liu, J. Qiao, W. Lin, and W. Zhang, "Air pollution monitoring by integrating local and global information in self-adaptive multiscale transform domain," *IEEE Transactions on Multimedia*, vol. 27, pp. 3716–3728, 2025, doi: 10.1109/TMM.2025.3535351.
[12]. L. Wang et al., "Causal-inference machine learning reveals the drivers of China's 2022 ozone rebound," *Environmental Science and Ecotechnology*, vol. 24, Art. no. 100524, Mar. 2025, doi: 10.1016/j.ese.2025.100524.
[13]. F. Ding et al., "Using multiple machine learning algorithms to optimize the water quality index model and their applicability," *Ecological Indicators*, vol. 172, Art. no. 113299, Mar. 2025, doi: 10.1016/j.ecolind.2025.113299.

[14]. Q. Aini, D. Manongga, U. Rahardja, I. Sembiring, and Y.-M. Li, "Technology readiness and behavioral intention to use air quality monitoring solutions—Measurement model," *PsycTESTS Dataset*, American Psychological Association (APA), 2025, doi: 10.1037/t96415-000.
[15]. N. Slimani, I. Slimani, N. Sbiti, and M. Amghar, "Machine learning and statistic predictive modeling for road traffic flow," *International Journal of Traffic and Transportation Management*, vol. 3, no. 1, pp. 17–24, Mar. 2021, doi: 10.5383/jttm.03.01.003.
[16]. K. S. Rautela, S. Singh, and M. K. Goyal, "Resilience to air pollution: A novel approach for detecting and predicting aerosol atmospheric rivers within Earth system boundaries," *Earth Systems and Environment*, vol. 9, no. 1, pp. 93–115, Jul. 2024, doi: 10.1007/s41748-024-00421-0.

ISSN (Online): 3105-9082