International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Detecting Environmental Anomalies: Variational Autoencoder- Based Analysis of Air Quality Time Series Data

Elmenawy Osman [1], Chitreshh Banerjee *[2], Ajeet Poonia [3], Prasad Sankarapu [4]

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: Air pollution presents a major environmental challenge, necessitating effective methods for rapid detection of pollution episodes to protect public health and economic interests. This study proposes a novel method combining Variational Autoencoders (VAEs) and Random Forest classifiers to identify anomalies in multivariate air quality time series data. The analysis focuses on key pollutants (NO2, PM10, PM2.5, O3, CO, and SO2) and meteorological variables, utilizing data from three monitoring stations over three years. By applying pre-processing techniques and dataset balancing with SMOTE, the hybrid model's performance is evaluated using various metrics. The results highlight the model's robustness in detecting air quality anomalies across different scenarios. Moreover, t-SNE visualizations of the encoded latent space reveal discernible patterns. This study underscores the potential of integrating deep learning with ensemble learning to improve air quality monitoring systems and suggests avenues for future enhancements in broader environmental monitoring applications.

Keywords: *Anomaly Detection, Variational Autoencoder, Unsupervised Learning, Air Quality*

1. Introduction

The escalating urbanization in Indian cities exacerbates poor air quality, negatively impacting residents' well-being [1]. The exceeding of National Ambient Air Quality Standard (NAAQS) limits by pollutants like particulate matter (PM10 and PM2.5) underscores the importance of air quality monitoring [2]. Monitoring primarily relies on networks of devices, often employing low-cost sensors within wireless sensor networks supported by Internet of Things (IoT) technology [3]. However, limitations in these sensors, such as variability and drift, challenge the reliability of air quality data [4].

Multivariate time series data, crucial for monitoring, faces challenges of missing values post-anomaly removal, affecting predictive models' performance [5]. Consistent monitoring aids in identifying abnormalities, crucial for various systems' integrity [6]. Anomaly detection is vital for averting unexpected issues across diverse systems [7].

Traditional methods and machine learning approaches may struggle with complex data, while Variational Autoencoders (VAEs) offer promise by modeling complex data distributions [8].

[1] *Amity Institute of Information Technology, Amity University, Jaipur, Rajasthan, India, minnawi22.osman@gmail.com , ORCID ID: 0009-0000-6227-9228*
[2] *Amity Institute of Information Technology, Amity University, Jaipur, Rajasthan, India,* chitreshh@gmail.com, *ORCID ID: 0000-0002-2303-9894*
[2] *University College of Engineering & Technology Dean (IIR), Bikaner, Rajasthan, India, pooniaji@gmail.com*
ORCID ID: 0009-0001-2079-9062
[3] *Harman Company, Principle Engineer, Hyderabad, India, spstats.aiml@gmail.com*
* Corresponding author's Email: chitreshh@gmail.com

Recent studies have integrated deep learning layers into VAEs for effective imputation, leveraging multivariate dependencies and dynamic characteristics of air quality data [9]. This study utilizes VAE to detect anomalies in air quality time series data collected from the Rajasthan State Pollution Control Board (RSPCB), India, via Continuous Ambient Air Quality Monitoring Stations (CAAQMS)

2. Literature Review

The literature review covers air quality monitoring and anomaly detection.

2.1. Air Quality Monitoring

Monitoring air quality is essential for urban areas due to the detrimental effects of pollution on health and the environment [10]. Traditional monitoring techniques often lack the spatial resolution to capture localized pollution hotspots [11]. Low-cost sensors and machine learning algorithms have been employed to enhance monitoring accuracy [12] [13]. Efforts have focused on optimizing calibration intervals and utilizing IoT systems for precise monitoring [14].

2.2. Anomaly Detection

Anomalies in air quality data pose significant risks, necessitating effective detection methods. Anomaly detection, widely applied across various fields, aims to identify outliers in data [15]. Anomaly detection can be conducted in batch or iterative modes, categorizing anomalies as point, contextual, or collective [16]. Time-series data may exhibit seasonal patterns, requiring specialized analysis approaches [17].

Machine learning-based anomaly detection models offer promising solutions but face challenges in multivariate data and spatial correlations [18]. Deep learning methods have shown remarkable progress in addressing these challenges, particularly in feature extraction and multivariate time-series analysis [19].

3. Background

The background section explores air quality time series data characteristics, common anomalies, and the application of Variational Autoencoders (VAEs) for anomaly detection.

3.1. Air Quality Time Series Data

The study utilizes real data from Continuous Ambient Air Quality Monitoring Stations (CAAQMS) in Jaipur, India, spanning from July 2017 to July 2020 (Rajasthan State Pollution Control Board,2020). This dataset encompasses key air quality measures (PM10, PM2.5, NO2, SO2, CO, O3) and meteorological factors with hourly temporal resolution. Data cleaning and preprocessing were conducted to address errors and missing values [20].

3.2. Common Anomalies in Air Quality Data

Anomalies in air quality data, such as spikes, sudden changes, and outliers, are vital for understanding environmental dynamics and potential hazards. Seasonal variations, equipment failures, and data gaps contribute to challenges in anomaly detection [21] Data visualization aids in identifying anomalies, but existing approaches have limitations, including computational expense and the need for extensive domain knowledge [22].

3.3. Deep Methods for Anomaly Detection

Deep learning techniques have advanced significantly, particularly in capturing complex time series patterns. Various deep anomaly detection models, including forecasting and reconstruction-based approaches, outperform traditional methods in real-world applications [23]. VAEs, introduced in 2014, excel in modeling intricate data distributions and are increasingly used for anomaly detection tasks in time series data [24]

3.4. Variational Autoencoders (VAEs)

VAEs leverage probabilistic latent spaces to capture intricate patterns and produce normal data distributions during training. By calculating reconstruction loss, VAEs determine anomaly scores, making them adept at identifying deviations from typical behavior. Their unsupervised learning capability enables anomaly detection in scenarios with limited labeled data [25]. The study applies VAEs to detect anomalies in air quality time series data, leveraging their capabilities to enhance anomaly detection in dynamic datasets.

4. Methodology

In the methodology section, we detail the VAE architecture utilized for anomaly detection, focusing on training processes and parameter fine-tuning.

4.1. VAE Architecture

The VAE consists of an encoder and decoder network trained jointly to balance reconstruction accuracy and encoding compactness [26]. The encoder employs convolution operations to map input data to a lower-dimensional latent space representation, while the decoder reconstructs data from the sampled latent space using polynomial trend blocks and transposed convolution [27].

Anomalies can be addressed by VAEs' probabilistic approach, as they predict mean and standard deviation parameters in a latent space Gaussian distribution. The encoder outputs variance and mean parameters of a multivariate normal distribution in the latent space, and during sampling, a random noise term is introduced to produce the latent vector. The decoder reconstructs input data from the sampled latent space, and a loss function, comprising reconstruction and regularization terms, guides VAE training. The combined loss function minimization ensures realistic output generation and meaningful latent space representations. This probabilistic approach allows VAEs to handle variability in the latent space and achieve robust anomaly detection [16, 28]. Fig 1 illustrates the structure of the VAE, comprising an encoder network, a decoder network, and a latent space sampling mechanism. A detailed explanation of each component is presented subsequently.

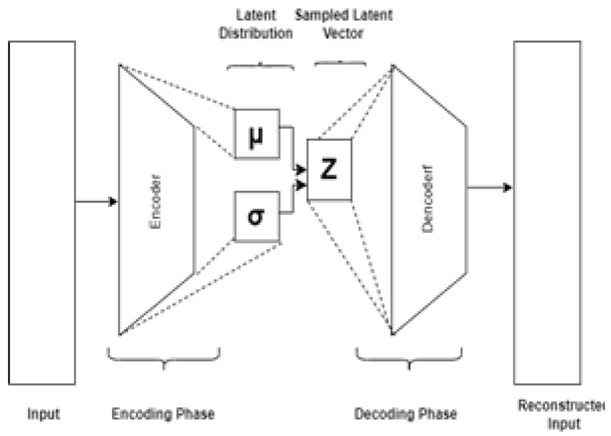


Fig. 1 illustrates the simple Structure of VAE

4.1.1. Encoder Network

An input data point is transformed by the encoder into a distribution inside the latent space. Typically, it consists of several layers of convolutional or fully connected neural network units, which progressively reduce the dimensionality of the input. The output of the encoder represents the variance and mean parameters of a multivariate normal distribution in the latent space [29].

4.1.2. Latent Space Sampling

Probabilistic dynamics are at play in the latent space of a VAE, where input data is not directly mapped to a specific point but rather encoded as a probability distribution. The encoder yields both the mean (μ) and log variance (log (σ^2)) of a multivariate normal distribution [30] .During sampling a random noise term (ϵ), drawn from a standard normal distribution, is introduced to the mean, the resulting latent vector is then calculated as

μ + ϵ. Exp (0.5 * log (σ^2))
(1)

Introducing stochasticity in the training process. This sampling mechanism aligns with the Eq. (2)

$$z = \mu\varphi(x) + \sigma\varphi(x) \odot \varepsilon \quad (2)$$

Where

Z: represents the sampled latent vector.

μ: the mean vector outputted by the encoder network.

Φ(x): refers to the encoded representation of the input data x by the encoder.

σ : The standard deviation vector outputted by the encoder

⊙: denotes the element-wise multiplication

ε: A vector of random samples from a standard normal distribution

4.1.3. Decoder Network

The input data is reconstructed by the decoder using a sample taken from the latent space. It usually consists of multiple layers of neural network units that gradually increase the dimensionality back to the initial input size, much like the encoder [31]. The reconstructed data is represented by the decoder's output.

4.1.4. Loss functions

A particular loss function, comprising a reconstruction loss and a regularization term formula (2), is utilized by the VAE. The difference between the input and reconstructed data is measured by the reconstruction loss. Binary cross-entropy and mean squared error are popular options. The regularization term pushes the latent space to adhere to a particular structure (typically a standard normal distribution), and is frequently based on the Kullback-Leibler (KL) divergence [32]. This guarantees a continuous and smooth latent space and helps avoid overfitting.

Loss function = reconstruction loss + KL divergence loss (3)

4.1.5. Training

The above combined loss function, which takes into account the regularization term and the reconstruction loss, is minimized by the VAE during training. The decoder learns to produce realistic outputs from sampled latent vectors, and the latent space learns to encode meaningful representations of the input data .The inclusion of a probabilistic approach in the latent space sets VAEs apart from conventional auto encoders, enabling them to produce a wide range of distinctive and meaningful samples.

5. Experimental Setup

This section gives a thorough description of the dataset used in this research, including details about the data's provenance and the attributes that were examined. It describes the process for dividing the dataset into separate sets for testing, validation, and training. This section also lists the different evaluation metrics that were used to determine the anomaly detection model's effectiveness, providing a clear framework for performance evaluation.

5.1. Data description

The experimental dataset, sourced from the Rajasthan State Pollution Control Board (RSPCB) in India, was collected via Continuous Ambient Air Quality Monitoring Stations (CAAQMS) stationed in Jaipur. These stations, situated at various locations, automatically monitor air pollutants and meteorological data to provide a comprehensive view of air quality. The dataset spans from July 1, 2017, to July 1, 2020, comprising measurements of PM10, PM2.5, O3, CO, SO2, and NO2, alongside meteorological attributes. Preprocessing steps, including outlier detection using the Interquartile Range (IQR) method and addressing missing values through Nearest Neighbor Interpolation, were crucial to enhance data quality for subsequent anomaly detection with a Variational Autoencoder (VAE). Scaling the data between 0 and 1 using normalization concluded the preprocessing phase, ensuring suitability for deep learning models and improving overall model reliability and accuracy.

5.2. Implementation Details

The research utilized Spyder IDE version 5.4.3 and Python 3.11.5, operating within the TensorFlow framework version 2.15.0. Essential libraries including Scikit-Learn, NumPy, and pandas were also employed. Moreover, to balance the datasets, the Synthetic Minority Over-sampling Technique (STOME) technique was applied within TensorFlow. The project was carried out in the Anaconda 3 (2023) environment, using Python 3.11.5 (64-bit), with Qt 5.15.2, PyQt5 version 5.15.7, on a Windows 10 platform.

5.3. Data splitting for training, validation, and testing

The study established a cutoff date ('2019-11-24 13:00:00') to divide the dataset, allocating 80% for training and 10% for testing purposes. During training, the model adjusts its weights and biases to reduce the difference between its predictions and the actual observations. This training phase is crucial for the model to identify patterns, with the Variational Autoencoder (VAE) being utilized to reconstruct data, thereby improving its generalization capabilities. To mitigate the risk of overfitting—where the model performs well on the training data but poorly on unseen data—an additional 10% of the data was reserved for validation. This strategy ensures the model's robustness and effectiveness in handling new data.

5.4. Evaluation Metrics

Three performance metrics—precision, recall, and F1-score—that are frequently used assessment metrics in anomaly detection were employed in the comparison studies. Equations. (3) and (4) define the terms precision and recall

$$\text{Precision} = \frac{TP}{(TP+FP)} \tag{4}$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \tag{5}$$

Where FP denotes false positives, FN denotes false negatives, and TP stands for true positives. Where FP denotes the number of normal points that are falsely predicted as abnormal points, FN denotes the number of abnormal points that are predicted to be normal, and TP is the number of abnormal points that are correctly detected.

Accuracy represents the proportion of correctly predicted observations to the total predictions for a specific class, serving as a metric to assess the predictive quality of a model. Recall is defined as the fraction of correctly identified predictions out of all actual instances of the same category. A higher recall indicates that the model is more adept at identifying anomalies, making it a crucial metric. Precision, often used alongside recall to gauge model performance, can at times present conflicting results. Hence, for a more holistic evaluation of anomaly detection capabilities, the F1-score is considered. The F1-score, being the harmonic mean of precision and recall, offers a balanced measure of a model's accuracy and recall, providing a comprehensive view of its effectiveness. Equation (5) provides the definition of an F1-score.

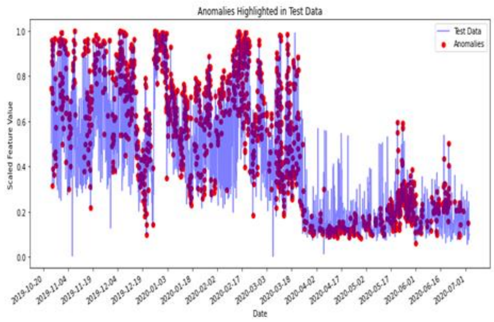
$$\text{F1} = 2 * \frac{(P*R)}{(R*P)} \tag{6}$$

The F1-score value can be calculated by taking the harmonic mean of the precision and recall rates. This allows for the simultaneous consideration of the model's accuracy and recall rates during the detection process. The formula uses P to stand for the detection model's accuracy rate and R for the detection's recall rate.

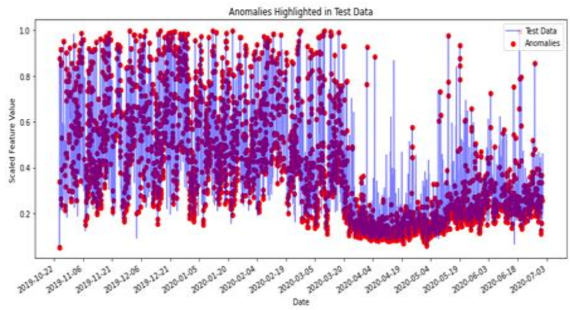
Area under the curve (AUC) (receiver operating characteristics (ROC)): Measures the capacity of the model to differentiate the classes [33] . The total percentage of items correctly classified is known as accuracy. It is computed by dividing the total number of predictions by the number of correct predictions, making it possibly the most straightforward metric.

6. Result & Discussion

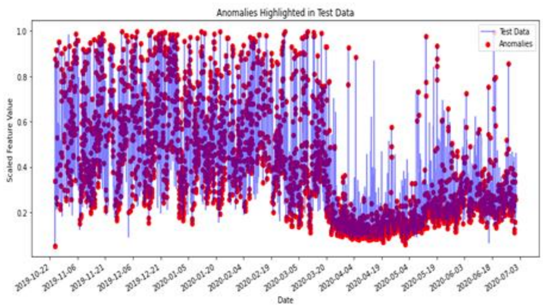
This study employs Variational Autoencoders (VAE) on three years of air quality data from three stations. It focuses on six pollutants and additional meteorological variables. Synthetic Minority Over-sampling Technique (SMOTE) balances class imbalance, enhancing anomaly detection. SMOTE, combined with Random Forest classifier, and improves model performance, validated by precision, recall, and F1-score metrics. This systematic approach ensures the classifier is trained on a balanced dataset and optimized for detecting anomalies. SMOTE application after threshold determination is crucial for addressing class imbalance, significantly enhancing anomaly detection in air quality data. Anomaly pinpointing in NO2 levels shows varying frequencies across monitoring areas. Figures 2 (a, b and c) focus on pinpointing anomalies in NO2 levels



(a)



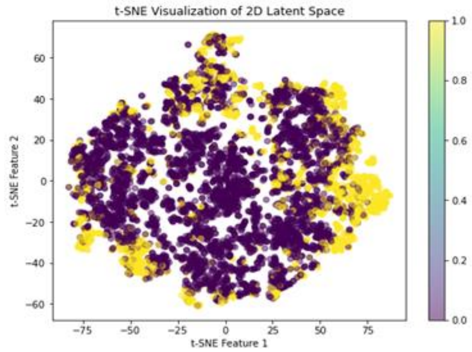
(b)



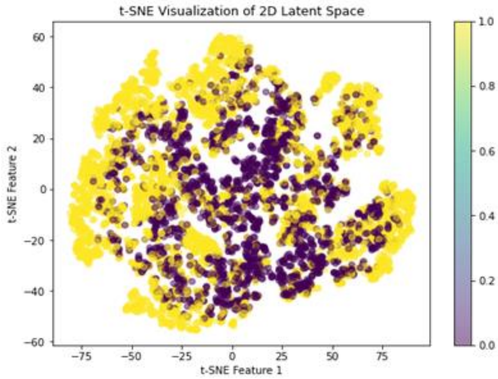
(c)

Fig. 2 Anomaly Detection in NO2 Levels of (a) Police Comm (b) Science Park And (c) Psy_Center Datasets

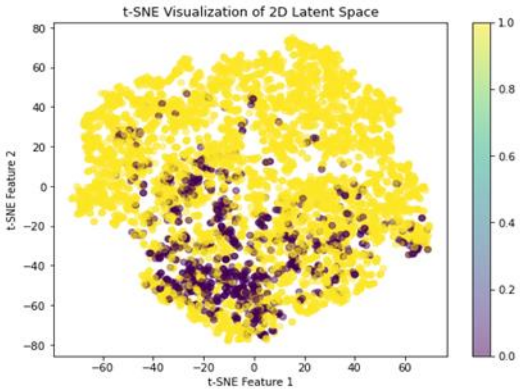
We utilized t-SNE to visualize latent spaces of datasets reduced by VAE, highlighting anomaly detection effectiveness and dataset-specific patterns. Figures 3(a,b and c) present the t-SNE visualizations for each of the three datasets.



(a)



(b)

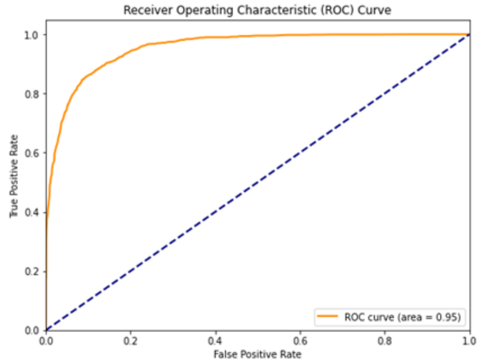


(c)

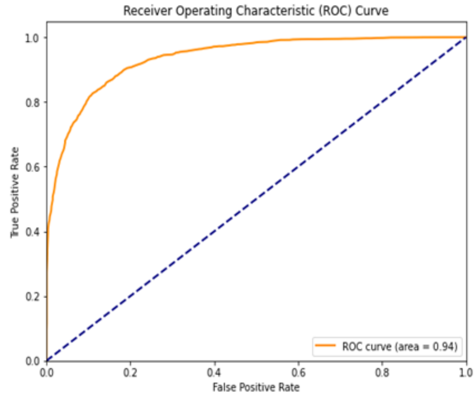
Fig. 3 t-SNE 2D latent space of of (a) Police Comm (b) Science Park And (c) Psy_Center Datasets

The results demonstrate VAE's effectiveness with SMOTE and Random Forest in detecting anomalies. Police Comm Dataset achieved 90.0% accuracy with high precision and recall. Science Park Dataset showed balanced precision and recall, with 86.27% accuracy. Psy_Center Dataset exhibited superior precision and recall, achieving 91.0% accuracy. The evaluation metrics for the model's performance across all test datasets are summarized in **Table 1.**

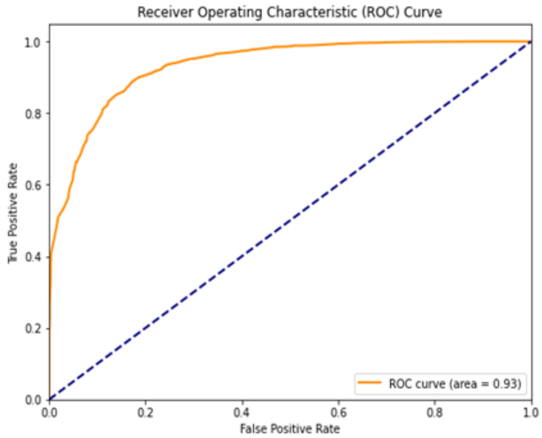
The suggested method for detecting anomalies examined by utilizing ROC curves and the area under the curve (AUC). The following graphs displays the ROC for the three stations: Police Comm, Science Park, and Psy_Center.



(a)



(b)



(c)

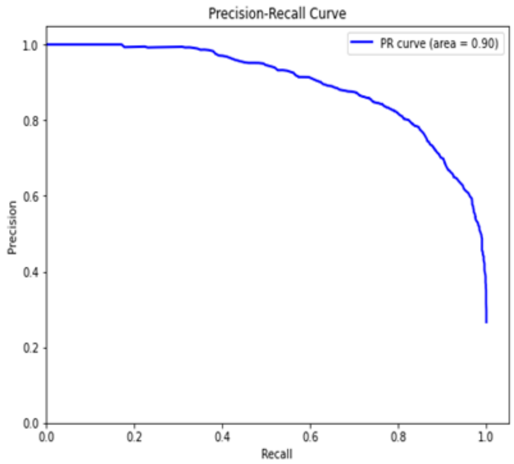
Fig.4 ROC curve for (a) Police Comm (b) Science Park And (c) Psy_Center Datasets

The ROC-AUC results reveal the VAE's robust anomaly detection across datasets. Specifically, the AUC for the Police Comm dataset reached 0.95, as depicted in Figure 4(a). Likewise, Figure 4(b) illustrates the Science Park dataset's AUC at 0.94, while Figure 4 **(c)** indicates the Psy_Center dataset's AUC at 0.93. These metrics underscore the VAE's efficacy, especially evident in the

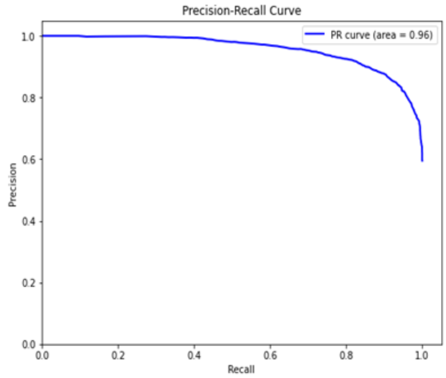
Police Comm dataset, where it outperformed others marginally. Furthermore, a deeper dive into anomaly detection involves examining classification thresholds, false positive rates (FPR), and true positive rates (TPR). Notably, the Police Comm dataset's ROC curve exhibits a balanced trade-off between TPR and FPR, maintaining FPR below 0.077 while achieving TPR exceeding 0.82 across various thresholds. This behavior signifies a well-calibrated model adept at distinguishing anomalies from normal data points.

Similarly, ROC curves for Park and Psy_Center stations demonstrate a balanced TPR and FPR, showcasing the model's accuracy in detecting anomalies at different thresholds. For instance, the Park dataset boasts an 85% TPR and 0.14 FPR, indicating efficient anomaly detection with a reasonable false alarm rate. Similarly, the Psy_Center dataset showcases a 95% TPR and 0.3 FPR, implying effective detection of genuine anomalies with minimal false positives.

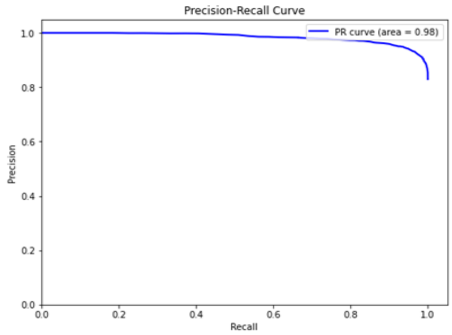
Additionally, the precision-recall curve's area under the curve (PR AUC) serves as a quantitative measure of the model's anomaly detection across datasets. High PR AUC values of 90%, 96%, and 98% for Police Comm, Science Park, and Psy_Center datasets, respectively at **Figures 5 (a,b and c)** reflect the model's ability to detect anomalies accurately while minimizing false positives, validating its resilience and flexibility in diverse environmental conditions.



(a)



(b)



(c)

Fig .5 RP curve for (a) Police Comm (b) Science Park And (c) Psy_Center Datasets

7. Challenges and Limitations

In the study on air quality anomaly detection, challenges emerged, emphasizing areas for further research. Ensuring data quality and completeness was paramount. Selecting relevant features amid environmental data complexity posed another hurdle. Imbalanced data between normal instances and anomalies hindered model training and evaluation. Resource-intensive advanced models prolonged computations. Determining optimal anomaly detection thresholds was complex, affecting model performance. Future improvements entail testing on diverse datasets, adjusting hyper parameters, reducing computational demands, and employing advanced feature engineering. Broadening evaluation metrics and enabling continuous learning would enhance models' adaptability and effectiveness in environmental monitoring and public health applications

Table 1 The evaluation metrics for the model's performance across all test dataset

Anomaly Detection								
Model	Datasets	Anomaly Detection Metrics						
		Precision		Recall		F1		Accuracy
		Class 0	Class 1	Class 0	Class 1	Class 0	Class 1	
VAE	Police Comm	0.94	0.80	0.92	0.83	0.93	0.81	90.0%

	Science Park	0.81	0.90	0.86	0.86	0.84	0.88	86.27%
	Psy_Center	0.75	0.94	0.70	0.95	0.72	0.95	91.0%

8. Conclusion and Future Work

This study presents a robust anomaly detection framework for air quality monitoring, showcasing the efficacy of VAEs, Random Forest classifiers, and SMOTE. Future research could explore alternative deep learning architectures like CNNs and RNNs for better spatial-temporal pattern capture. Incorporating more environmental variables may enhance model accuracy. Real-time anomaly detection implementation promises proactive environmental hazard responses. Practical application could lead to automated, intelligent air quality monitoring systems. These findings offer a foundation for improving anomaly detection across sectors like finance and healthcare, underscoring the model's adaptability and efficiency in diverse contexts.

Conflicts of Interest

The authors assert that they do not have any identifiable conflicting financial interests or personal relationships that could have been perceived to impact the work described in this paper.

Author Contributions

Conceptualization, Sankarapu and Osman ; methodology, Osman; software, Sankarapuand and Osman]; validation, Banerjee and Osman , formal analysis, Osman; investigation, Banerjee; resources, Sankarapu and Osman ; data curation, Osman; writing—original draft preparation, Osman ; writing—review and editing, Banerjee and Sankarapu; visualization, Osman; supervision, Banerjee and Poonia; project administration, Banerjee and Poonia

References

[1] C. S. Bhosale, et al., "Ambient Air Quality Monitoring with Reference to Particulate Matter (PM10) in Kolhapur City," Nature Environment and Pollution Technology, vol. 22, no. 4, pp. 2029-2037, 2023.

[2] S. K. Guttikunda and B. R. Gurjar, "Role of meteorology in seasonality of air pollution in megacity Delhi, India," Environmental Monitoring and Assessment, vol. 184, pp. 3199-3211, 2012.

[3] W. Jiao, et al., "Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States," Atmospheric Measurement Techniques, vol. 9, no. 11, pp. 5281-5292, 2016.

[4] G. Tancev, "Relevance of drift components and unit-to-unit variability in the predictive maintenance of low-cost electrochemical sensor systems in air quality monitoring," Sensors, vol. 21, no. 9, p. 3298, 2021.

[5] T.-B. Ottosen and P. Kumar, "Outlier detection and gap filling methodologies for low-cost air quality measurements," Environmental Science: Processes & Impacts, vol. 21, no. 4, pp. 701-713, 2019.

[6] F. Rollo, C. Bachechi, and L. Po, "Anomaly detection and repairing for improving air quality monitoring," Sensors, vol. 23, no. 2, p. 640, 2023.

[7] M. Fahim and A. Sillitti, "Anomaly detection, analysis and prediction techniques in IoT environment: A systematic literature review," IEEE Access, vol. 7, pp. 81664-81681, 2019.

[8] D. Park, Y. Hoshi, and C. C. Kemp, "A multimodal anomaly detector for robot-assisted feeding using an LSTM-based variational autoencoder," IEEE Robotics and Automation Letters, vol. 3, no. 3, pp. 1544-1551, 2018.

[9] J. Loy-Benitez, S. Heo, and C. Yoo, "Imputing missing indoor air quality data via variational convolutional autoencoders: Implications for ventilation management of subway metro systems," Building and Environment, vol. 182, p. 107135, 2020.

[10] T. Burki, "WHO introduces ambitious new air quality guidelines," The Lancet, vol. 398, no. 10306, p. 1117, 2021.

[11] J. S. Apte, et al., "High-resolution air pollution mapping with Google street view cars: exploiting big data," Environmental Science & Technology, vol. 51, no. 12, pp. 6999-7008, 2017.

[12] L. Spinelle, et al., "Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO2," Sensors and Actuators B: Chemical, vol. 238, pp. 706-715, 2017.

[13] N. Zimmerman, et al., "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring," Atmospheric Measurement Techniques, vol. 11, no. 1, pp. 291-313, 2018.

[14] M. Zaidan, et al., "Dense air quality sensor networks: Validation, analysis, and benefits," IEEE Sensors Journal, vol. 22, no. 23, pp. 23507-23520, 2022.

[15] [15] G. Li and J. J. Jung, "Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges," Information Fusion, vol. 91, pp. 93-102, 2023.

[16] X. Teng, Y.-R. Lin, and X. Wen, "Anomaly detection in dynamic networks using multi-view time-series hypersphere learning," in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 827-836, 2017.

[17] Q. Wang, et al., "A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting US shale oil production," Energy, vol. 165, pp. 1320-1331, 2018.

[18] N. Shaadan, et al., "Anomaly detection and assessment of PM10 functional data at several locations in the Klang Valley, Malaysia," Atmospheric Pollution Research, vol. 6, no. 2, pp. 365-375, 2015.

[19] S. Schmidl, et al., "Anomaly detection in time series: a comprehensive evaluation," Proceedings of the VLDB Endowment, vol. 15, no. 9, pp. 1779-1797, 2022.

[20] D. Liu, et al., "Aqeyes: visual analytics for anomaly detection and examination of air quality data," arXiv preprint arXiv:2103.12910, 2021.

[21] B. Actkinson and R. Griffin, "Detecting plumes in mobile air quality monitoring time series with density-based spatial clustering of applications with noise," Atmospheric Measurement Techniques Discussions, pp. 1-22, 2023.

[22] X. Cheng, et al., "Climate modulation of Niño3.4 SST-anomalies on air quality change in southern China: Application to seasonal forecast of haze pollution," Atmospheric Research, vol. 225, pp. 157-164, 2019.

[23] Y. Wei, et al., "LSTM-autoencoder-based anomaly detection for indoor air quality time-series data," IEEE Sensors Journal, vol. 23, no. 4, pp. 3787-3800, 2023.

[24] X. Shu, et al., "Unsupervised dam anomaly detection with spatial–temporal variational autoencoder," Structural Health Monitoring, vol. 22, no. 1, pp. 39-55, 2023.

[25] Z. Li, Y. Sun, L. Yang, Z. Zhao, and X. Chen, "Unsupervised machine anomaly detection using autoencoder and temporal convolutional network," IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-13, 2022.

[26] Desai, C. Freeman, Z. Wang, and I. Beaver, "TimeVAE: A variational auto-encoder for multivariate time series generation," arXiv preprint arXiv:2111.08095, 2021.

[27] N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, "N-BEATS: Neural basis expansion analysis for interpretable time series forecasting," arXiv preprint arXiv:1905.10437, 2019.

[28] O. Rybkin, K. Daniilidis, and S. Levine, "Simple and effective VAE training with calibrated decoders," in International Conference on Machine Learning, pp. 9179-9189, PMLR, 2021.

I. Higgins, L. Matthey, A. Pal, C. P. Burgess, X. Glorot, M. M. Botvinick, S. Mohamed, and A. Lerchner, "beta-VAE: Learning basic visual concepts with a constrained variational framework," ICLR (Poster), vol. 3, 2017.

A. J. Lew and M. J. Buehler, "Encoding and exploring latent design space of optimal material structures via a VAE-LSTM model," Forces in Mechanics, vol. 5, p. 100054, 2021.

[29] M. Memarzadeh, B. Matthews, and I. Avrekh, "Unsupervised anomaly detection in flight data using convolutional variational auto-encoder," Aerospace, vol. 7, no. 8, p. 115, 2020.

[30] S. K. Zhou and R. Chellappa, "From sample similarity to ensemble similarity: Probabilistic distance measures in reproducing kernel Hilbert space," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 6, pp. 917-929, 2006.

[31] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in Proceedings of the 23rd International Conference on Machine Learning, pp. 233-240, 2006.