frontiers | Research Topics

Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities

Edited by
Ketan Kotecha, Sayali Apte, Sayali Sandbhor and Amit Srivastava

Coordinated by
Vaishnavi Dabir

Published in
Frontiers in Environmental Science
Frontiers in Environmental Engineering
Frontiers in Sustainable Cities

ISSN 1664-8714
ISBN 978-2-8325-6374-8
DOI 10.3389/978-2-8325-6374-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities

Topic editors
Ketan Kotecha — Symbiosis International University, India
Sayali Apte — Symbiosis International University, India
Sayali Sandbhor — Symbiosis International University, India
Amit Srivastava — University of Adelaide, Australia

Topic coordinator
Vaishnavi Dabir — Symbiosis International University, India

Citation
Kotecha, K., Apte, S., Sandbhor, S., Srivastava, A., Dabir, V., eds. (2025). *Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6374-8

Table of contents

frontiers | Frontiers in Environmental Engineering

TYPE Editorial
PUBLISHED 09 May 2025
DOI 10.3389/fenve.2025.1612210

OPEN ACCESS

EDITED AND REVIEWED BY
Christian Kennes,
University of A Coruña, Spain

*CORRESPONDENCE
Sayali Apte,
sayali.apte@sitpune.edu.in

RECEIVED 15 April 2025
ACCEPTED 21 April 2025
PUBLISHED 09 May 2025

CITATION
Sandbhor S, Kotecha K, Apte S, Srivastava A and Dabir V (2025) Editorial: Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities.
Front. Environ. Eng. 4:1612210.
doi: 10.3389/fenve.2025.1612210

Editorial: Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities

Sayali Sandbhor[1], Ketan Kotecha[2], Sayali Apte[1]*, Amit Srivastava[3] and Vaishnavi Dabir[4]

[1]Department of Civil Engineering, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Pune, India, [2]Symbiosis Centre for Applied Artificial Intelligence (SCAAI), Dean of Engineering, Symbiosis International (Deemed University), Pune, India, [3]Faculty of Sciences, Engineering and Technology, University of Adelaide, Adelaide, SA, Australia, [4]Green Cube Consulting LLC, Georgia, United States

KEYWORDS

environment, artificial intelligence, smart and sustainable cities, interdisciplinary, sustainable development

Editorial on the Research Topic
Artificial intelligence in environmental engineering and ecology: towards smart and sustainable cities

Introduction

The escalating global population intensifies resource scarcity, biodiversity decline, and climate change, placing immense pressure on ecosystems and generating numerous social, economic, and environmental crises. Embracing sustainable development presents a pathway to mitigate these environmental challenges through enhanced resource efficiency and conservation efforts. Given the intricate relationship between population growth and its environmental ramifications, a collaborative interdisciplinary strategy, integrating diverse fields of knowledge, is essential. Recognizing the Sustainable Development Goals as a blueprint for a viable future, interdisciplinary research within civil and environmental engineering assumes paramount importance. It is now critical to formulate interdisciplinary solutions aimed at realizing these goals, particularly focusing on fostering sustainable cities and communities, ensuring access to affordable and clean energy, providing clean water and sanitation, promoting responsible consumption and production patterns, advancing industry, innovation, and infrastructure, and enacting effective climate action. Meeting the pressing demands for sustainable urban environments, accessible and clean energy sources, safe water and sanitation systems, conscientious consumption and production practices, advancements in industry, innovation, and infrastructure, and robust climate action requires the immediate development of interdisciplinary solutions to achieve the Sustainable Development Goals.

This Research Topic "Artificial Intelligence in Environmental Engineering and Ecology: Towards Smart and Sustainable Cities" is floated with an intention to capture the essence of latest research to advance the field of environment and sustainability. This compilation

showcases cutting-edge innovations and actionable solutions aimed at achieving the Sustainable Development Goals, particularly highlighting interdisciplinary engineering research pertinent to SDGs 3, 6, 7, 9, 11, 12, 13, 14, and 15 for a sustainable future. The purpose of this Research Topic is to bring together a range of innovative studies pushing the boundaries of thought on integrating environmental engineering and artificial intelligence. The editorial outlines the contributions and the ways in which they help us better understand application of AI in environmental engineering for smart cities. With contributions from researchers across several countries including China, United States, India, Turkey, Iran, Denmark, Republic of Korea, Romania, Australia, the Research Topic features 10 scholarly papers that include three review articles and seven original research papers. These ten papers in this Research Topic, delve into critical Research Topic such as water contamination, prediction of air pollutants, pile fire detection, reducing smart city corporate carbon intensity and overall environmental pollution. The researchers employ a variety of AI methodologies including deep learning, Convolutional Neural Networks (CNN), Gene Expression Programming (GEP), Graph Attention Networks (GAT), K-Nearest Neighbor (KNN) Fuzzy Modeling, Causal Spatio-Temporal Attention Networks (Causal-STAN), and autoencoders (AE).

Review articles: establishing the role of artificial intelligence in environmental and infrastructure monitoring

The three review articles present the potential of utilizing AI capabilities in environment and smart city related problems to arrive at a suitable solution. Kaveh and Alhajj summarize various approaches to crack detection to enhance the current practices, Chakurkar et al., present potential of AI in detecting cracks in infrastructure works in smart cities. This review highlights two key approaches for the detection of cracks that are deep learning and traditional computer vision. In addition to the crack detection, the review also discusses ways to quantify the crack severity level. Similarly, Popescu et al., through a review, have summarised AI solutions for hazardous substance monitoring in different environments including air, water and soil along with AI powered technologies for pollution monitoring including spectroscopy, ground-based monitoring sensors, aerial imaging and unmanned aerial vehicles (UAVs), ground robotics, satellite remote sensing.

In addition to the review articles the Research Topic comprises of following original researches.

Original research: innovative applications of AI in environmental engineering

The seven original research papers present leading AI techniques applied to a range of environmental challenges. Each contributes unique insights and solutions aimed at achieving sustainable development goals through smart technologies.

Chao and Qiu present a novel fuzzy modeling approach for estimating air pollution concentrations that integrates an enhanced evidence theory with comprehensive weighting and the K-nearest neighbor (KNN) interval distance within the framework of the matter-element extension model. Capable of reducing the error rate by 38% relative to alternative methods, the authors point towards improving computational efficiency using AI in air quality monitoring.

Similarly, Bashardoost et al., present potential of autoencoder (AE) in denoising the input data for deep learning and then using it for building CNN algorithm to carry out spatial modeling and risk mapping of air pollutants including PM 2.5, PM 10, SO_2, NO_2, O_3, and CO. With prediction accuracies ranging from 0.8 to 0.96, the study highlights CNN-AE model's impressive precision when generating the pollution risk map. Meanwhile, Wang et al. explore GEP to construct the relationship between pollutant gas and $PM_{2.5}/PM_{10}$ and back propagation neural networks (BPNN) as a baseline method. With the experimentation, they demonstrate usability of the proposed approach in environmental modelling. Wang et al. develop spatio-temporal attention causal convolutional neural network (Causal-STAN) architecture for predicting PM2.5 concentrations, proving it to be a superior architecture as compared to recurrent network model.

Furthermore, Joshi et al. present investigations of the application of multiple machine learning algorithms to detect wildfires. Applying novel deep learning with ensemble approach to the early detection has proven to be promising. Anaadumba et al., propose use of Graph Attention Network (GAT) to predict lead contamination in drinking water proving to be better than XGBoost machine learning algorithm. Peng et al. demonstrate using the difference-in-differences model to empirically analyze the impact of the smart city pilot policy on corporate carbon intensity.

Key studies demonstrate significant advancements in prediction accuracy, risk mapping, and early detection of environmental hazards. The Research Topic underscores the growing role of AI in tackling air and water pollution and calls for more research in areas like soil contamination and solid waste management. Contributions from countries across multiple continents demonstrate the global interest in applying AI to sustainability challenges. The Research Topic has benefited from a broad international authorship, with China contributing the most papers. The use of diverse AI tools from traditional models to cutting-edge neural networks shows that tailored solutions are often effective and necessary depending on the specific problem being addressed. Use of AI methods could help address more comprehensive environmental goals promising a more resilient, efficient, and sustainable future.

The Research Topic is guest-edited by scholars from Australia and India, with editorial coordination led by Dr. Vaishnavi Dabir. The editorial team wish to extend gratitude to the entire Frontiers team for their continuous support throughout the publication process. This collaborative effort has offered a platform for researchers to share their perspectives and innovations in applying AI to environmental engineering.

Author contributions

SS: Writing – original draft. KK: Writing – review and editing. SA: Writing – review and editing. AS: Writing – review and editing. VD: Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

Author VD was employed by Green Cube Consulting LLC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Sustainable Cities

TYPE Review
PUBLISHED 25 October 2023
DOI 10.3389/frsc.2023.1253627

OPEN ACCESS

EDITED BY
Gresha Bhatia,
Vivekanand Education Society's Institute of Technology, India

REVIEWED BY
Mehdi Gheisari,
Islamic Azad University, Iran
Madhumita Takalkar,
Crown Institute of Higher Education (CIHE), Australia
Vaishali Khairnar,
Terna Engineering College, India
Debahuti Mishra,
Siksha 'O' Anusandhan University, India

*CORRESPONDENCE
Deepali Vora
✉ deepali.vora@sitpune.edu.in

RECEIVED 05 July 2023
ACCEPTED 06 October 2023
PUBLISHED 25 October 2023

CITATION
Chakurkar PS, Vora D, Patil S, Mishra S and Kotecha K (2023) Data-driven approach for AI-based crack detection: techniques, challenges, and future scope.
Front. Sustain. Cities 5:1253627.
doi: 10.3389/frsc.2023.1253627

Data-driven approach for AI-based crack detection: techniques, challenges, and future scope

Priti S. Chakurkar[1,2], Deepali Vora[3]*, Shruti Patil[4], Sashikala Mishra[3] and Ketan Kotecha[3]

[1]Computer Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India, [2]School of Computer Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India, [3]Department of Computer Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India, [4]Department of Artificial Intelligence and Machine Learning (AIML), Symbiosis Institute of Technology, Symbiosis International (Deemed University) (SIU), Pune, Maharashtra, India

This article provides a systematic literature review on the application of artificial intelligence (AI) technology for detecting cracks in civil infrastructure, which is a critical issue affecting the performance and longevity of these structures. Traditional crack detection methods involve manual inspection, which is laborious and time-consuming, especially in urban areas. Therefore, automatic crack detection with AI technology has gained popularity due to its ability to identify degradation of roads in real-time, leading to increased safety and reliability. This review emphasizes two key approaches for crack detection: deep learning and traditional computer vision, with a focus on data-driven aspects that rely primarily on data from training datasets to detect and quantify the severity level of the crack. The article highlights the advantages and drawbacks of each approach and provides an overview of various crack detection models, feature extraction techniques, datasets, potential issues, and future directions. The research concludes that deep learning-based methods used for crack classification, localization and segmentation have shown better performance than traditional computer vision techniques, especially in terms of accuracy. However, deep learning methods require large amounts of training data and computational power, which can be a significant limitation. Additionally, the article identifies a lack of 3D datasets, unsupervised learning algorithms are rarely used to train crack detection model, and datasets having road images with variety of road textures such as asphalt and cement etc. as challenges for future research in this field. A need for 3D and combined texture datasets as challenges for future research in this field.

KEYWORDS

computer vision, crack detection, deep learning, image processing, segmentation

1. Introduction

Detecting cracks in vital infrastructure such as roads, bridges, and buildings cost millions of rupees annually. Catching flaws on highways and roads has also been a topic of attention for health safety to ensure driving safety. Natural calamities, such as foods and earthquakes, cause significant damage existing infrastructure. Because cracks are the most visible and widespread expression of civil work structure deterioration, developing operative methods

for monitoring and treating cracks is critical to maintaining the structural health of civil engineers (Li et al., 2020; Wang et al., 2021). One of the biggest problems in the world today is road safety. Given how frequently it is used, maintaining decent road pavement is essential to reducing accidents, and consequently, the number of fatalities. Overloading, seepage, inadequate and poor road surface drains, absence of appropriate road maintenance, a lack of proper design, and unsuitable climatic conditions, among other things, are significant reasons for rutting and degradation. Road distresses like cracking and disintegration obstruct and have a detrimental effect on traffic flow and safety, resulting in poor road performance. Early detection of road cracks is crucial to take adequate corrective actions before the issue gets out of hand and the pavement worsens. Maintenance processes typically involve a visible examination and estimate of the current state to maintain damaged infrastructure's structural and functional integrity. These damages could appear as minor or significant cracks that worsen over time, eventually leading to the structure's collapse or destruction. Cracks appear in many infrastructures (Cao M. T. et al., 2020; Munawar et al., 2021a; Wang et al., 2021) (tunnels, bridges, roads, pipelines, etc.) throughout their useful life that can reveal and enhance possible structural pathologies. Therefore, the detection of these cracks is vital in inspection work.

In most cases, the inspection is visual and performed by persons. This manual inspection is laborious, expensive, and longer duration. Due to the required knowledge, expertise, and experience, as well as human mistake brought on by fatigue and inattention, it also has a limited level of dependability, impartiality, and reproducibility. An accurate and efficient alternative to the human procedure is to use image recognition to monitor engineering structures and algorithms for machine learning to understand the photos and extract the crucial geometric details of the fracture. Automatic crack detection is essential to safeguard infrastructure's effectiveness and durability. Figure 1 depicts the real-time applications where computer-based crack detection is needed Chen Y. et al. (2021) suggested machine learning methods in surface defect detection is a key part in the quality inspection of industrial products. First, according to the use of surface features, the application of traditional machine vision surface defect detection methods in industrial product surface defect detection is summarized from three aspects: texture features, color features, and shape features. Secondly, the research status of industrial product surface defect detection based on deep learning technology from three aspects: supervised method, unsupervised method, and weak supervised method. Ali et al. (2021) studied surface cracks on the concrete structures as a key indicator of structural safety and degradation. To ensure the structural health and reliability of the buildings, frequent structure inspection and monitoring for surface cracks is important. Surface inspection conducted by humans is time-consuming and may produce inconsistent results due to the inspectors' varied empirical knowledge. The employment of deep learning algorithms using low-power computational devices for a hassle-free monitoring of civil structures.

Crack detection with digital image processing is the essential step toward automation in road health monitoring. Research and business have been discreetly moving toward developing and applying computerized road surface monitoring systems to reduce expenses associated with manual inspection. An automated system for road crack detection must be built in four steps: acquisition of the image, pre-processing of an image, image segmentation, crack detection, and classification (Mohan and Poobal, 2018). Each of these steps has its importance in the system.

Artificial Intelligence (AI) based technology offers a more sophisticated approach for crack detection, which can execute various tasks (such as classification or regression) with exceptional performance. AI-based Crack features can be extracted using hand-crafted feature engineering with computer vision and automatic feature extraction with a deep learning approach—AI-based crack detection categories as computer vision-based crack detection and deep learning-based crack detection. Morphological operations, edge detection algorithms, support vector machine (SVM), and random structured forests (RSF) are the classical approaches used in the literature to extract hand-crafted features from asphalt pavement/road images. Wang et al. (2019) analyzed five classification algorithms, the support vector machines (SVM), neural networks (NN), random forests (RF), logistic regression (LR), and boosted tree (BT), to classify rail surface cracks. However, the practical application of this method is limited due to its slow convergence, over-fitting, and high computational cost, etc. Therefore, a fast and automatic deep learning-based feature extraction algorithm with CNN, transfer learning, and general adversarial network (GAN) (Zhong K. et al., 2020) are used to process the considerable monitoring data. Ghaderzadeh et al. (2022) proposed fully automatic deep learning based system for Acute Lymphoblastic Leukemia (ALL) diagnosis (ALL) diagnosis and subtype classification (early pre-B, pro-B, and pre-B ALL). The system's overall procedure would entail feeding the network pairs of segmented and original images, feature extraction using DenseNet-201, and then use the classification block to predict the ALL subtype based on the retrieved features. The model learns the intricate correlations and patterns between the input data and their corresponding subtypes through training on a sizable dataset of labeled images. Ghaderzadeh et al. (2021) suggested a deep learning-based model for COVID-19 identification using X-ray image. This model has the potential to enhance current testing strategies and aid in the pandemic response. Hosseini et al. (2023) study shows the proposed mobile application's potential to be an effective screening tool for hematologists and clinical professionals. The application can accurately detect B-ALL cases by utilizing preprocessing methods and deep learning algorithms, which can assist eliminate needless bone marrow biopsy cases and shorten the time required for B-ALL diagnosis. Gheisari et al. (2023) highlighted how deep learning (DL) techniques can be selected based on various issues and uses. It draws attention to the possibility for future work in developing DL frameworks and investigating novel applications, like forecasting natural disasters.

1.1. Significance

The importance of AI-based automatic crack detection systems is rising expeditiously hand in hand with sensor technology and the internet. Image sensors are economically helpful compared to

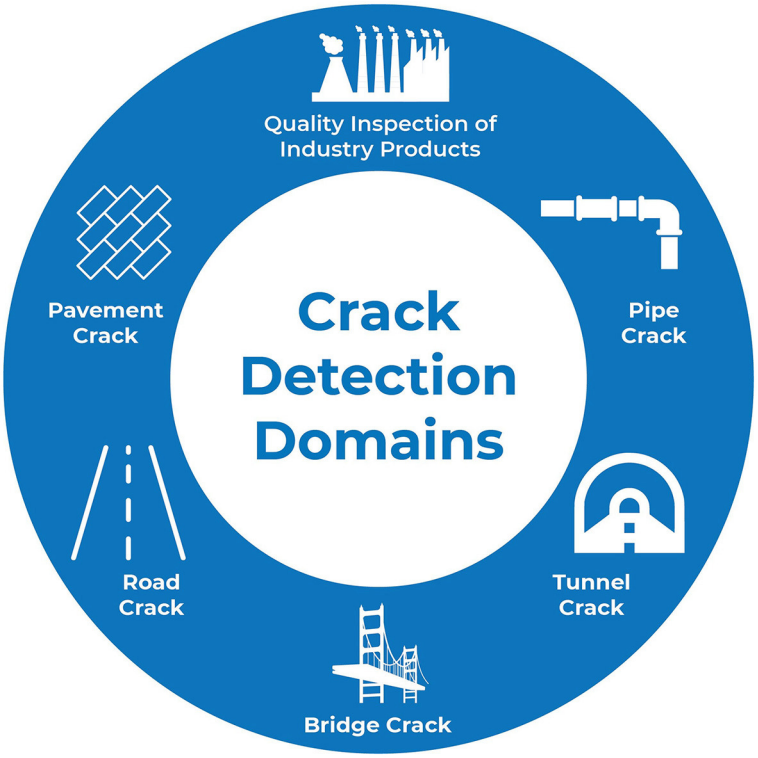


FIGURE 1
Crack detection application domains.

other sensors, which capture real-time images of civil structures. Crack features extracted from the images can help the timely and proactive management of structures. Research and business have been discreetly moving toward developing and applying AI-based crack detection systems to reduce expenses associated with manual inspection and ensure greater safety.

1.2. Motivation

There have been many reviews of the literature on image-based crack detection systems. Still, to our knowledge, challenges and probable solutions for Computer Vision-based and Deep Learning based crack detection system have yet to be the subject of a systematic literature review. Therefore, it is necessary to consider the systematic reviews of methodologies, datasets, and evaluation metrics used. This review article presents the challenges and future scope of the AI-based crack detection system, which guides researchers in constructing more effective and reliable crack detection systems in the future.

1.3. Prior research

Crack detection with image processing is a research area with many literature review articles. To our best knowledge, a systematic literature review has yet to be written on the data-driven approach of AI-based crack detection. Indeed, AI models' performance is evaluated based on the quality of the input data. With a data-driven perspective, we studied how data is vital to improving the performance of AI-based crack detection. Hence, to achieve reliability and efficiency in crack detection, it becomes sensible to consider systematic reviews on the data-driven approach in this survey. Table 1 overviews survey papers studied in crack detection with image processing.

The study by Nguyen et al. (2022) evaluates the effectiveness of deep learning-based crack detection algorithms in locating cracks in asphalt pavement. This study recommends using pix2pix for crack segmentation, ResNet and DenseNet for crack classification, and Faster R-CNN for crack object detection. However, they suggest further research on unsupervised and semi-supervised learning techniques to improve fracture identification in asphalt pavement.

In a different article, Hamishebahar et al. (2022) provide a comprehensive literature review of deep learning-based crack detection research and evaluate studies that utilize the same publicly accessible datasets to determine effective crack detection strategies. Based on the trends and evaluated papers, the report suggests important avenues for future research in crack detection. Cao M. T. et al. (2020) provide a comprehensive overview of techniques for detecting road pavement cracks, including image processing, machine learning, and 3D imaging. The article compares and discusses deep learning neural networks for crack detection based on classification, object detection, and segmentation approaches, highlighting their significant improvement in detection performance. The study also evaluates

TABLE 1 Summary of existing surveys on crack detection.

Ref.	Data set	Data collection	Preprocessing methods	AI based models			Learning			Performance metric	Challenges	Future direction
				Crack classification	Crack Object detection	Crack segmentation	Supervised	Semi/Weakly supervised	Unsupervised			
Nguyen et al. (2022)	X	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	PathGANs metric is a new approach with high-grade results GPR images can be used to identify cracks inside the pavement in depth Supervised learning GANs should be used to solve crack segmentation. Semi supervised and unsupervised learning should be focus on
chch Hamishebahar et al. (2022)	✓	X	X	✓	✓	✓	X	X	X	X	✓	Deep edge detection crack segmentation task in unsupervised way, Complexity of the crack shapes in depth and the parameter correlations with the performances of the different proposed architectures and loss functions in the area.
Cao M. T. et al. (2020)	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	Comparison and analysis on deep learning methods and 3D image based methods Collecting 3D crack benchmark datasets will greatly benefit future study of the 3D crack detection.
König et al. (2022)	✓	X	X	✓	✓	✓	✓	✓	✓	✓	✓	Semi, weakly and unsupervised domain in crack detection will push the research forward and alleviate the issues that come with small datasets and difficult annotations.

(Continued)

10.3389/frsc.2023.1253627

TABLE 1 (Continued)

Ref.	Data set	Data collection	Preprocessing methods	AI based models			Learning			Performance metric	Challenges	Future direction
				Crack classification	Crack Object detection	Crack segmentation	Supervised	Semi/Weakly supervised	Unsupervised			
Li et al. (2022c)	✓	X	X	✓	✓	X	X	✓	✓	✓	✓	Analyzed crack detection through the pixel-level CIS methods to identify the improvement directions of the future work to better use DL method to enhance the performance and efficiency of CIS
Hsieh and Tsai (2020)	X	✓	X	✓	✓	✓	X	X	X	✓	✓	ML crack detection algorithms, reviewing 68 ML-based crack detection papers to identify the current development trend, pixel level crack segmentation. Performance comparisons among 8 DL crack segmentation models were then conducted using consistent evaluation metrics and real-world 3D pavement images under diverse conditions
Munawar et al. (2021b)	✓	✓	X	✓	X	X	X	X	X	X	X	Need to perform severity assessment of the detected cracks

10.3389/frsc.2023.1253627

the performance of these approaches on widely-used benchmark datasets and covers performance evaluation measures.

Golding et al. (2022) propose using convolutional neural networks (CNN) as a deep learning-based technique for fracture detection in infrastructure and comparing grayscale and RGB models using various image processing methods. The results indicate that DL crack identification does not rely on color, as grayscale models perform similarly to RGB models, thresholding, and edge detection models perform worse than RGB models. König et al. (2022) discuss the importance of early surface crack detection and monitoring for structural health monitoring and provide a review of deep learning-based crack analysis algorithms. The study covers a range of tasks, including crack classification, detection, segmentation, and quantification, and offers thorough analyses of current fully, semi-, and unsupervised techniques. The review also includes measures used to assess algorithm performance and well-known datasets used for cracking.

Hu et al. (2021) discuss using deep learning models to identify asphalt pavement cracks, highlighting the issues with conventional artificial detection systems. They show that pavement fracture detection using the YOLOv5 series deep learning model has produced positive results, with the YOLOv5l model having the best detection accuracy (88.1%) and the YOLOv5s model having the quickest detection time (11.1 ms per image). In the study published by Li et al. (2022a) examine the importance of fracture detection in transportation infrastructure and the growth of deep learning-based techniques for crack image segmentation (CIS). They conduct a thorough analysis of over 40 papers on DL-based CIS methods released in the previous three years, categorizing them into ten themes based on backbone network design, including FCN, U-Net, multi-scale, attention mechanism, transformer, and weakly supervised learning, among others.

In this study, Hsieh and Tsai (2020) provide a thorough analysis of recent machine learning-based crack detection algorithms, with a special focus on pixel-level crack segmentation. They evaluate eight ML-based models using standardized evaluation criteria and 3D pavement photos with various conditions, showing that deeper backbone networks and skip connections improve performance in FCN models. The suggested algorithm tackles the false-positive issue as a necessary first step to enhance ML-based crack detection models.

1.4. Research goals

This systematic literature review (SLR) aims to understand recent developments in the field of computer vision-based road crack detection techniques and identify unresolved problems and obstacles within. The AI-based crack detection system survey's research goals are listed below.

- What are the different artificial intelligence-based approaches used for crack detection?
- What are the different datasets available for research purposes in AI-based crack detection?
- How to accurately measure the segmented crack parameters to assess the cracks' severity?
- What performance evaluation indicators are employed to assess the effectiveness of the AI-based crack-detecting system?
- How to optimize AI-based crack detection models be implemented with limited computational resources?

1.5. Work's contribution

The following are the critical findings of this in-depth literature review:

- To conduct a literature study integrating deep learning and computer vision-based crack detection systems.
- To provide a depth summary of the existing research on deep learning and computer vision-based crack detection systems with an emphasis on methods, data sets, applications, challenges, and future directions.
- We also provide an overview of publicly available datasets that can support research in the crack detection domain.
- To evaluate this method's potential applications for automatic road health monitoring, as well as its current advantages and disadvantages.

This article is a comprehensive and well-structured study on image-based road crack detection, covering various aspects such as a literature review on data-driven AI approaches, computer vision techniques, challenges, and future scope.

2. Research strategy

Published research is accessible based on core metrics used in the bibliometric analysis. As part of this process, we attempt to select the most well-known or active researchers and their affiliations, collaborative patterns, frequently used phrases, and numerous articles about them. We termed the initial step of our research process as "image-based crack detection." "The phrase crack-detection" is used in civil and structural sectors such as tunnels, bridges, and railway tracks for structural health monitoring.

2.1. Database selection and query formulation

This research aims at finding the use of artificial intelligence in the image-based crack detection domain; we started with bibliometric analysis for image-based crack detection as a primary keyword ANDing with "Computer Vision," "Deep Learning," "Machine Learning," Artificial Intelligence," as a secondary keyword in the research space. As a result, these primary and secondary keywords are used to frame the query, and relevant articles were initially gathered from the Scopus database. The primary and secondary keywords used in questions to select the data from Scopus databases using AND and OR Boolean operators are shown below.

> TITLE-ABS-KEY (crack detection) and ((computer vision) or (deep learning) or (machine learning) or (artificial intelligence)) and(image).

After that, the filtration technique is used for the article collection to enhance the outcomes that satisfy our main objectives. This procedure's steps are removing duplication, applying exclusion and inclusion criteria, filtering based on titles and abstracts, and full-text screening. Inclusion criteria used to filter the relevant documents, such as

- Publication Articles must be published between the years 2012 to 2022.
- The subject area should be Computer science, decision sciences, material sciences, and multidisciplinary articles.
- The article should match a minimum of one of the search terms.
- The article type should be conference papers or review articles.

Exclusion criteria are foreign language research articles and articles unrelated to research questions.

2.2. Analysis of the information

In this review article, bibliometric analysis was done on filtered reports from the Scopus database using various parameters, including:

- Subject area
- Research trends in AI-based technologies
- Documents by country
- Keyword co-occurrences.

2.2.1. Subject area

To begin with, an analysis of the data revealed that the keyword "image-based crack detection" is most frequently utilized within the domain of Computer Science, followed by Engineering and Material Science. Material science is essential to the field of structural health monitoring (SHM) because of how much it affects a structure's performance, dependability, and durability. Since image-based crack detection is primarily a computer vision problem, it is unsurprising to observe the highest concentration of research activity within the Computer Science subject. This trend is visually represented in Figure 2.

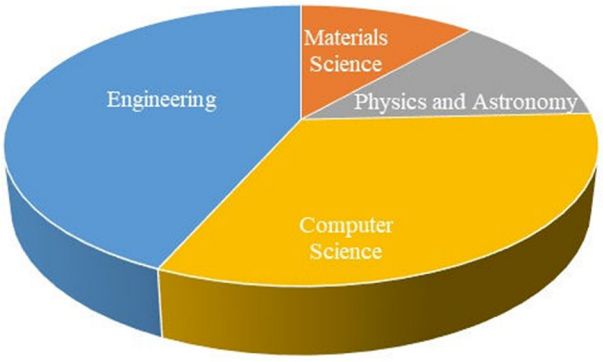


FIGURE 2
Analysis of AI-based crack detection research in various subject areas.

2.2.2. Research trends in AI-based technologies

As shown in Figure 3, the graph demonstrates the trends that focus on the number of research articles retrieved from the Scopus database with primary keyword image-based crack detection with secondary keywords such as computer vision (CV), machine learning (ML), deep learning (DL), and artificial intelligence (AI). The number of research papers focusing on crack detection using CV keywords increased from below 50 in 2018 to over 150 in 2022, with more than 50 research papers already published in 2023. Similarly, AI keyword usage increased from around 20 research papers to approximately 50 in 2022, with about 30 published in 2023. ML keyword usage rose from 30 research papers in 2018 to over 100 in 2022, and around 60 were published in 2023. Notably, DL keyword usage increased from approximately 30 research papers in 2018 to 350 in 2022, and it has been used in over 160 research papers in 2023. These trends indicate a growing interest in and utilization of CV, AI, ML, and DL techniques in crack detection research.

2.2.3. Documents by country

The graph shown in Figure 4 represents a country-wise distribution of research papers focusing on AI-based crack detection. Among the countries mentioned, Hong Kong is associated with around 100 research papers, India with approximately 400 research papers, and notably, China with over 2600 research papers. The graph indicates a significant research output in AI-based crack detection from these countries. While the specific numbers for other countries are shown above, it is evident that multiple nations, including Italy, Germany, Japan, Australia, Canada, the United Kingdom, South Korea, and the United States, are actively engaged in this field. However, China has emerged as a dominant player with many research papers, showcasing its strong presence and advancing AI-based crack detection.

2.2.4. Keyword co-occurrences

Analysis of co-occurrence is performed on phrases using a VOS viewer. Keywords occurring with more than 50 are only considered for the analysis. 13,242 keywords out of 66 met the criteria. Keyword co-occurrence network visualization is shown in Figure 5 below and identifies terms that can be utilized more frequently in this research.

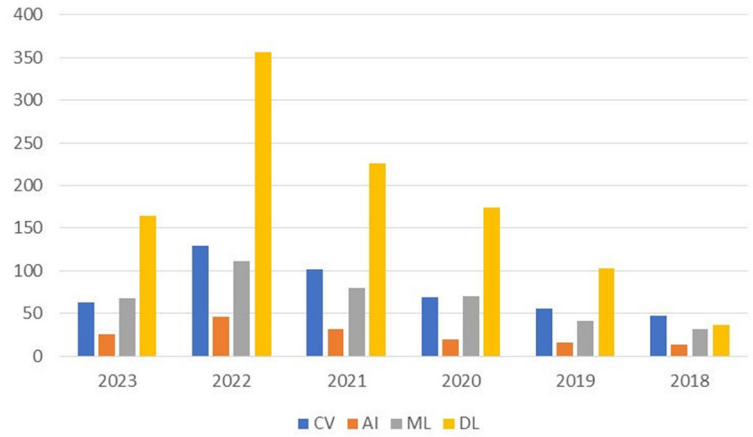


FIGURE 3
Research trends in image-based crack detection with different AI-based technologies.

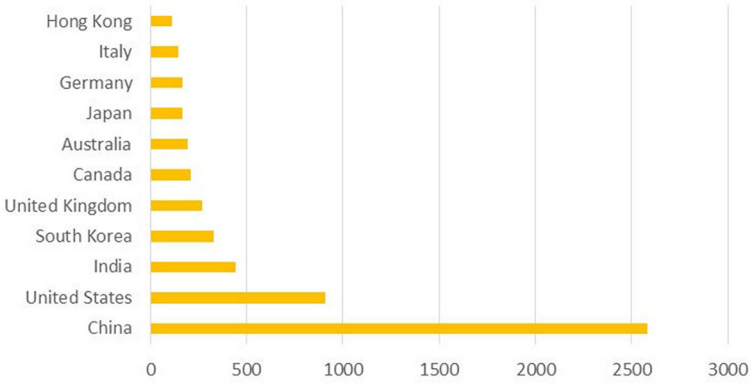


FIGURE 4
Country-wise distribution of research papers focusing on AI-based crack detection.

3. Data-driven AI-based imaging system for crack detection

Data is crucial in AI research, especially for image processing and crack detection tasks. Obtaining high-quality data is vital because the accuracy and effectiveness of AI models heavily rely on the data they are trained on Figure 6. Highlighted crack detection process from the data perspective, and the survey is presented in different data-oriented aspects used for crack detection.

Figure 6 shows an AI-based data-driven crack detection process. The process can be broken down into four stages, each of which is essential in detecting cracks. The first stage involves data collection, labeling, and dataset building. Data is collected using UAV-mounted cameras and handheld devices. The collected data is labeled at the pixel, object, and image levels, and private and public datasets are used. This stage is critical for ensuring that the data used for crack detection is accurate and reliable. Data augmentation, preprocessing, and learning algorithms are used in the second stage. Data augmentation is done using Generative Adversarial Networks (GANs), which help generate synthetic data and improve the system's accuracy. Preprocessing is done using Histogram Equalization Filters, which normalize the images and improve their contrast. Learning algorithms such as supervised, weakly supervised, and unsupervised are used to train the system. The third stage involves crack classification, object detection, and segmentation. Transfer learning techniques such as VGG and MobiNet are used to develop the crack classification model. Crack object detection uses techniques such as YOLO, FASTER RCNN, and SSD.

Crack segmentation is performed using FCN, DEEPV LAN, and encoder-decoder techniques. These techniques are used to identify and locate cracks in the images accurately. The fourth and final stage uses evaluation metrics, attention techniques, and crack severity qualification. Evaluation metrics such as IOU, MAP, and AUC are used to assess the system's performance. Attention techniques such as SENET, CBAM, ECANET, and COORD ATT are used to improve the system's accuracy. Finally, crack severity qualification is done by measuring the detected cracks' length, width, and depth. This information can then be used to prioritize repairs and maintenance.

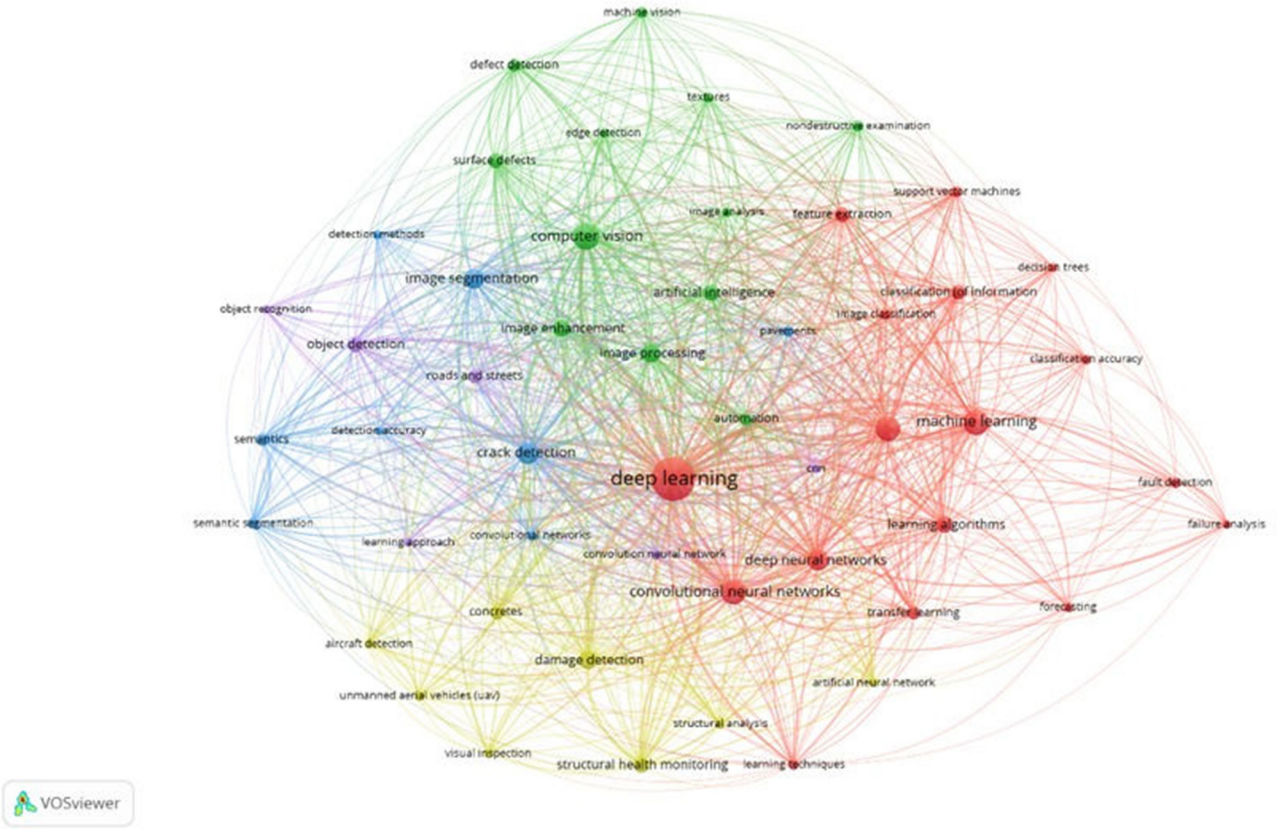


FIGURE 5
Network visualization diagram for keyword co-occurrence.

Overall, this approach leverages AI-based techniques for efficient and accurate crack detection, which can be applied in various industries and infrastructures, such as roads, bridges, and buildings, to ensure their safety and longevity.

4. Literature survey on data-driven AI-based imaging system for crack detection

The term "Image-based crack detection system" describes the full spectrum of activities, from taking images to classifying cracks according to their severity. This system provides the economic and engineering analysis tools required for making cost-effective maintenance, rehabilitation, and reconstruction decisions. A smart road monitoring system based on Industry 4.0 was developed. In addition to video, mobile data, weather data, and other sensor data, the monitoring system collects substantial amounts of data from the road environment infrastructure. Monitoring systems collect data on road environment defects that provide road environment safety data.

Consequently, road environment monitoring systems may include pavement and bridge crack detection. Internal, invisible faults and surface apparent defects include pavement and bridge defects. Surface obvious crack faults have been a long-standing issue that endangers public safety. The photographs of the road crack were taken in different lighting conditions (round the clock). A few of them included undesirable things, e.g., random particles, textures, heterogeneity, uneven lighting with variations in the road's surface, lines, boisterous environment, shadows, water, tire prints, oil slicks, etc. The result is a challenging problem in choosing a uniform threshold building. A successful preprocessing stage is crucial for getting good results because of the segmentation step. This action involves bringing edges, borders, or contrast to sharper focus analysis (Fu et al., 2020). There are numerous types of image enhancement like fuzzy edge removal, noise reduction, magnification, contrast stretching, filtering, artificial interpolation coloring, transform operations, histogram modeling, and fake color etc.

Automated image-based crack feature extraction methods can be divided into deep learning and traditional approach. The conventional systems mainly include the threshold approach, wavelet transform, the minimum path, etc. Traditional visual methods primarily rely on hand-craft features, with distinguishing abilities in the images, such as the grayscale, texture, and contour shape of defects, because of the road's intricacy, variety of topology, random forms, and sizes, as well as oil stains, weed stains, and other significant disruptions. Hand-craft features tend to fail, and the algorithm needs to be redesigned.

FIGURE 6
Overview of data-driven AI-based crack detection.

Therefore, traditional crack feature extraction methods have weak generalization ability and low efficiency for different road images in complex situations. Hence, research requirement is still active due to low robustness and fluctuating environment. Deep learning can resolve complex problems automatically with the help of AI. Various methods of detecting road cracks using deep learning have been classified according to their accuracy:

Patch level crack detection – likely to be binary or multi-crack classification.

Object Level-Likely to be boundary box regression.

Pixel-level crack detection - likely to be a semantic segmentation task

In recent studies, several methods for crack detection have been proposed using deep learning. The process may be based on the detection of objects or the segmentation of blocks of images. AI based crack detection process consists of four steps: data collection, data pre-processing, dataset modeling and crack classification as shown in Figure 7.

4.1. Data collection

The data collection is the initial stage that feeds the input images to the entire crack detection cycle. Automating data acquisition has led to the development of complete systems carried out by vehicles (or, more recently, on devices such as smartphones or crewless aerial vehicles) for visual surface surveying. A computer, a Global Positioning System (GPS) sensor, and imaging sensors are all installed in the car. Several sensing technologies, sensor positions, and vehicles have been deployed to collect data to evaluate surface conditions. Table 2 summarizes the sensor systems for data collection.

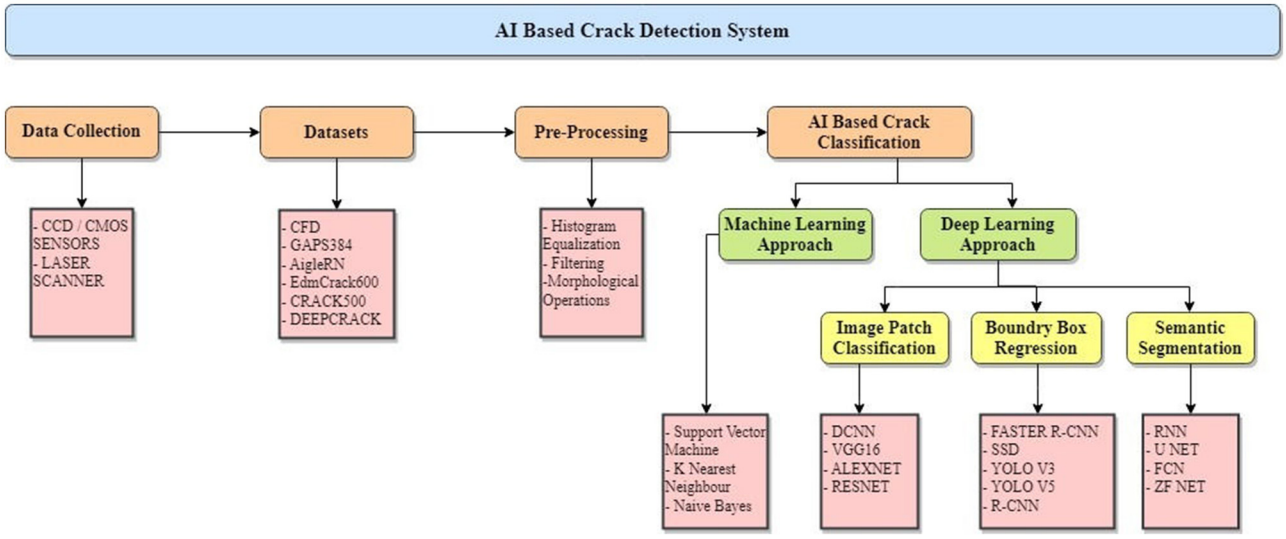


FIGURE 7
AI-based crack detection tasks and algorithms.

To ensure the safety, maintaining structural integrity, and lowering maintenance costs, cracks should be found and examined. Conventional techniques of fracture detection can be costly, time-consuming, and frequently necessitate stopping operations. However, the advent of AI-based data-driven crack detection systems, which can be quickly and precisely identify cracks, is a result of technological breakthroughs. Data collection, which uses sensor systems for fracture detection, is the basis of these systems. UAVs, camera-mounted vehicles, and handheld devices are just a few of the crack-detection sensor systems available for data collection. Each of these methods has its benefits and drawbacks, and choosing one depends on several considerations, including the level of accuracy that is sought, the difficulty of the terrain, and the available budget.

One of the most often used sensor systems for crack detection is crewless aerial vehicles (UAVs). They have cameras, sensors, and GPS and can take pictures at different angles and heights (Li et al., 2023). Depending on the type of camera and sensors being utilized, UAVs can capture both 2D and 3D images (Khaloo et al., 2018). The visual view can be in the top, front, or both directions at once. An ordinary UAV-based crack detection system typically consists of the UAV itself, a camera or sensor payload, a GPS, and a ground control station. A pilot controls the flight path and camera settings while controlling the UAV remotely. The application and required level of precision can influence the camera or sensor payload. High-resolution cameras and sensors are frequently utilized to take precise surface images for crack identification (Byrne et al., 2017). The UAV's location is tracked using GPS, ensuring reliable data collection. The ground control station tracks the UAV, gathers data, and performs data processing.

The capacity of UAVs to capture images from various angles and heights, which can give a more thorough perspective of the surface being investigated, is one benefit of employing them for data collection. UAVs can check substantial infrastructure projects like highways and bridges since they can swiftly and effectively cover large regions (Munawar et al., 2021a). However, the cost of purchasing and operating UAVs can be high, and their usage might be prohibited in some places due to laws or safety concerns.

Another alternative for gathering data on crack detection is a vehicle with surveillance cameras. Typically, they include a car with cameras or sensors that take pictures of the examined surface (Montero et al., 2015). Depending on the camera and sensors utilized, camera-mounted vehicles can record both 2D and 3D images. The top view, front view, or a combination of the two can be the image view. The car, a camera or sensor payload, and a computer system for data processing make up a conventional camera-mounted vehicle-based crack detection system. The driver, who controls the vehicle's direction and speed, is in charge of it. The application and required level of precision can influence the camera or sensor payload. High-resolution cameras and sensors are frequently utilized to take precise surface images for crack identification. The computer system does the data collection and processing. Data collecting with camera-mounted vehicles has the benefit of minimal energy usage and simple control. They are appropriate for checking both horizontal and vertical surfaces and can rapidly and effectively cover huge regions (Cao W. et al., 2020). The usage of camera-mounted vehicles, however, can be restricted in some situations, such as enclosed locations, and they might not be able to take pictures from specific angles or heights.

Another form of sensor system for crack detection uses handheld devices. An operator can use these portable tools to take pictures of surface cracks (Sony et al., 2019). Mobile devices with high-resolution cameras, such as smartphones and tablets, are the most often utilized handheld devices for crack detection. These tools can take pictures of the cracks in 2D and 3D (Chen et al., 2014). The portability and simplicity of use of mobile devices are advantages. They are a practical choice for small-scale crack detection applications because they are reasonably priced and easily accessible. The drawback of handheld equipment is that an operator is needed to take the images. This indicates that gathering data

TABLE 2 Sensor systems for data collection in road crack detection.

Crack detection sensor system	Components	Image type		Image view		Advantages	Limitations
		2D	3D	Top view	Front view		
UAV (Ersoz et al., 2017; Tse et al., 2023)	Cameras, sensors, GPS, etc. RGB images, IMU	✓	✓	✓	✓	Easy to capture images in dangerous and/or unreachable places	Images collected by UAV can be more challenging due to lighting conditions, viewpoints, and scales
Camera mounted vehicle (Mei and Gui, 2020; Yang N. et al., 2022)	Camera, GPS	✓	✓	X	✓	Low energy consumption and easy control	Difficult to capture images in dangerous situations
Handheld devices (Kong et al., 2017; Jeong et al., 2020)	Smartphones	✓	X	X	✓	Simple and low cost	Need to compromise with complete view, and image quality. Not practical for large-scale projects

can take a long time and may need to be more practical for large-scale projects. The accuracy of the crack identification algorithms may also need to be improved by the lower image quality of handheld devices compared to UAVs or vehicles with mounted cameras (Jordan et al., 2018). Terrestrial Laser Scanning (TLS) is used to assess road surface conditions. A laser line scanner, High-resolution, continuous 3D road surveys were produced by using a laser scanner or digital camera placed on a moving vehicle. Zhong M. et al. (2020) showed a multi-sensor, laser scanning approach for crack identification by merging to generate a 3D simulation of the highway, laser line scan data was combined with video that recognized fissures as small as a few centimeters in depth as well as their location.

Hassan et al. (2022) proposed three-dimensional reconstruction systems, the LCMS system (Laser Crack Measurement System), consisting of two three-dimensional laser profile-meters capable of measuring the cross-sectional profiles of a road with a resolution of 1 mm. This system collects intensity information, as well as geometric report of the road surface, which allows for characterizing and displaying images together with the shape (texture) of the road. The crack's depth, which is regarded as a crucial primitive, can be calculated with the use of the image's 3D representation. This feature is used to categorize crack types as small, moderate, or severe. With the advent of 3D road scanners, it is anticipated that, in the near future, a higher level of accuracy can be attained by utilizing the depth information from the collected 3D model of the road pavement. Since there are currently no benchmark datasets for 3D crack identification, collecting these datasets will be very helpful for future research.

Ultimately, data collection for AI-based crack detection systems relies on crack detection sensor systems. UAVs, vehicles with cameras placed on them, and handheld gadgets are the three most often employed sensor systems. Due to their swift ability to cover broad regions, UAVs are well-suited for large-scale projects, although they can be expensive to operate and require trained operators. Mid-scale projects can benefit from camera-mounted vehicles since they are convenient, inexpensive, and straightforward. Handheld devices offer a cost-effective alternative for small-scale projects, but they need an operator to take the pictures. The choice of which sensor system to utilize relies on the scope and complexity of the project. Each sensor system has advantages and limits. The project's size, complexity, and resource availability all affect the sensor system that will be used to collect data. AI-based crack detection technologies could revolutionize our ability to detect and maintain infrastructure, guaranteeing that our buildings, bridges, and roads are secure and safe for years to come.

4.2. Data labeling and dataset generation

Data labeling is the key to training the detection model, and the quality of labeling determines the accuracy. The three tracks of classification, detection, and segmentation are distinct and require different methods of data labeling. Image classification

annotation involves assigning a binary label to indicate the presence or absence of an object in the image at a high level. In pavement damage classification, a standard method is dividing a large image into smaller sub-images, called patch-level classification. Object detection annotation is done at the level of individual objects, requiring the identification of the object's category and its position in the image, typically represented by a bounding box. Segmentation annotation involves labeling each pixel in the image with its corresponding category or background and is done at a pixel-level granularity. Patch-level classification and object-level detection provide information about the localization of objects. Table 3 summarizes the datasets specified in the literature for crack detection with the level of annotations used. For data labeling, the labeling tools were studied in the literature to manually create image labels to produce a crack dataset in PASCAL VOC data format. The labeling tool will generate XML files based on the rectangular boxes, including crack category, size, and location. We should avoid other objects entering the rectangular box during the data labeling, leading to a high false detection rate. For many years, road crack detection systems have been a study issue. There are several public datasets available to assist us in conducting improved research. Shi et al. (2016) this study advances the field of road cracking detection through presenting a new annotated dataset (Crack Forest Dataset) and an effective detection technique that surpasses competing options in terms of noise reduction and processing time. The study's emphasis on two evaluation indicators, the "Continuity index (CI)" and "Crack detection accuracy" to assess accuracy, ensures an accurate assessment of the suggested method's efficacy. Liu et al. (2019) study makes a substantial contribution to crack detection and segmentation by presenting the DeepCrack benchmark dataset, a ground-breaking public benchmark aimed at promoting standardized evaluation and method comparison. The dataset includes several, thoroughly hand-annotated crack images in a variety of scales and settings. In contrast to earlier studies, which relied on limited, narrow evaluation datasets, the new benchmark database, which consists of 537 RGB color photos with manual annotations. Arya et al. (2021) proposed RDD2020 dataset to facilitate the development and improvement of deep learning algorithms in order to advance automatic road damage identification. It stimulates academics to look into cutting-edge methods for precisely identifying and locating road damage. The RDD2020 dataset is rigorously built to contain a varied collection of incidents of road damage, recorded across various road types, conditions, and environmental factors. Bounding boxes that exactly define sections of road damage are carefully inserted to each image in the dataset. For the sake of testing and refining deep learning algorithms, these annotations serve as the ground truth.

4.3. Data preprocessing in crack detection

Road image preprocessing is essential in automated pavement crack detection using image processing and deep learning algorithms. This process involves various techniques that aim to enhance image quality, reduce noise, and improve the visibility of cracks in the image. One of the primary steps is converting the colored image into grayscale, simplifying the image analysis process. Histogram equalization balances the brightness and contrast of the grayscale image, making cracks more visible to algorithms. Filters like Gaussian, Median, and Sobel are commonly applied to remove noise and highlight edges, making it easier for algorithms to detect cracks. Morphological transformation using techniques like erosion, dilation, opening, and closing helps to remove unwanted objects and enhance essential features. The ultimate goal of road image preprocessing is to improve the accuracy of automated pavement crack detection systems by highlighting the elements of interest, such as cracks, in the input image. The process can significantly improve the efficiency of assessing road conditions and ensuring safety.

Yang Z. et al. (2022) have developed a method for pavement crack localization and segmentation algorithms that aims to automate the detection of cracks on roads, improve safety, and reduce maintenance costs. The proposed technique involves a three-stage algorithm based on digital image processing and deep learning, which includes a two-dimensional discrete wavelet transform to obtain low and high-frequency coefficients of the wavelet, Retinex algorithm, guided filtering based on wavelet transform, and soft threshold filtering for denoising. The ultimate goal of this approach is to enhance the accuracy and efficiency of pavement crack detection and segmentation algorithms. On the other hand, Chen et al. (2022) proposed a crack detection method using image processing techniques to improve road safety and reduce maintenance costs. The method involves a potential crack region method that uses multiple thresholds for crack detection, in which histogram equalization is used to adjust the grayscale value distribution and enhance local contrast for better distinction between the crack and background. The technique also uses mean filtering to remove noise and improve image quality, which aims to improve the accuracy and efficiency of crack detection on road surfaces.

In the study, Li et al. (2022a) propose a new technique that combines convolutional neural networks and hybrid image processing to improve the accuracy of crack classification and segmentation on concrete bridge images. To achieve this, they use a bilateral filtering technique to sharpen the crack details and highlight their characteristics while minimizing the influence of factors such as rain or stains. The processed images are then converted to grayscale and subjected to contrast enhancement to enhance visual performance. This proposed technique aims to make the maintenance of concrete bridges more efficient and cost-effective. Similarly, Parrany and Mirzaei (2022) offer an image processing strategy that uses the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm and non-linear diffusion filtering techniques to improve the accuracy and efficiency of surface crack identification in building structures, especially under non-uniform illumination conditions.

Automated crack segmentation aims to increase the precision and efficiency of detecting cracks in building structures. Nnolim (2020) proposes a method that uses saturation channel threshold, area classification, and a modified level set segmentation fusion with Canny edge detection. The studied article use global histogram equalization to normalize the distribution of intensities and the

TABLE 3 Summary of datasets used for crack detection.

Dataset	Basic features			Advanced features			Dataset	Basic features	Advanced features
	Image/ Sample no.	Resolution	Colored	Image/ Sample no.	Resolution	Environmental or other interfering factors			
Crack forest dataset (CFD)1	118	480 × 320	Crack Forest Dataset (CFD)1	118	480 × 320	Lane lines, shadows, noise like oil stain	For Non-commercial research purposes only ·	iPhone5	China
AigleRN2	38	991 × 462 311 × 462	AigleRN2	38	991 × 462 311 × 462	Pre-processed to reduce the non-uniform illumination	With more complex texture than CFD, ·Similar small data sets like ESAR and LCMS	Professional Camera	French
Crack5003	500	2000 × 1500	Crack5003	500	2000 × 1500	Lane lines, oil/wet stain, tire brake marks, speckle noise, etc.	·Covers almost all kinds of features except for shadow	LG-H345 Cell Phone	US
GAPs3	1969	1920 × 1080	GAPs3	1969	1920 × 1080	Little noise like oil stain, asphalt and rails, lane	–	JAI Pulnix TM2030 monochrome cameras, Kodak KAI2093	Germany
Cracktree2003	206	800 × 600	Cracktree2003	206	800 × 600	only shadow & lane lines	–	Unknown	–
G454	122	2048 × 1536	G454	122	2048 × 1536	Brightness change, oil stain, tire brake marks, speckle noise, etc.	·4 types: transverse, longitudinal, block, and alligator cracks	Unknown	China
EdmCrack6005	600	1920 × 1080	EdmCrack6005	600	1920 × 1080	Weather, illumination, shadow, texture difference	·Only for academic research ·Perspective image taken from the rear camera of a car	Rear Camera of Car – GoPro 7	Canada
JapanRoad6	9053	600 × 600	JapanRoad6	9053	600 × 600	Real street view perspective images captured from the front windshields of a car	·License: CC BY-SA 4.0 ·For road defect detection ·PASCAL VOC annotation format	Front windshields of Car LG-5X	Japan
SDNET20187	230 (Cropped to 56092)	4096x3840 (256 × 256)	SDNET20187	230 (Cropped to 56092)	4096x3840 (256 × 256)	with a variety of obstructions, including shadows, surface roughness, scaling, edges, holes, and background debris	·License: Creative Commons Attribution 4.0 ·54 bridge decks, 72 walls, 104 pavements ·Crack size: 0.06-25mm	16-MP Nikon camera	US
Concrete crack images for classification 8	458(cropped to 40000)	4032 × 3024 (227 × 227)	Concrete crack images for classification 8	458(cropped to 40000)	4032 × 3024 (227 × 227)	Shadows, lighting spot, blurred, including thin & close-up cracks	·Similar to SDNET2018, images collected from METU Campus Buildings	Professional Camera	Turkey

10.3389/frsc.2023.1253627

wavelet-based enhancement algorithm to enhance crack detection by amplifying detail coefficients, suppressing approximation coefficients, and increasing contrast and sharpness. This method, known as the modified HSI-based crack detection, aims to improve crack detection accuracy and efficiency. On the other hand, Yu et al. (2022) propose a deep learning-assisted image processing technique. The approach involves using a mask filter to remove handwritten marks and a ratio filter to eliminate speckle linear noise from the input images. The use of deep learning algorithm is to classify images, identify regions of interest, then segment the images using threshold techniques, and perform crack quantification using image features. The proposed method aims to enhance the accuracy and speed of crack detection and quantification in concrete bridges.

The study proposes an advanced deep-learning fusion model (Feng et al., 2020) for more precise pavement crack detection and segmentation. The process starts with applying median and bilateral filters to eliminate noise, followed by contrast enhancement to improve the quality of the images by highlighting crack information. The fusion model combines convolutional neural networks and fully connected networks to achieve better results. Its main goal is to enhance the efficiency and accuracy of pavement crack detection and segmentation algorithms.

4.3.1. Data augmentation

Data augmentation techniques can be highly beneficial when you have limited training data for tasks like artificial crack detection. Data augmentation involves applying various transformations or modifications to existing data to create additional training examples, effectively expanding the dataset without the need for new annotations. This process helps improve the model's performance by providing more diverse and varied examples. Data augmentation is a machine learning technique that artificially generates additional synthetic training data through label-preserving transformations (Bayer et al., 2023). Though data augmentation techniques have achieved great success in computer vision applications, Data augmentation is performed by rotating the images with different angles, adding some noise, etc. The CrackNet, a deep-learning framework for crack detection and classification leverages data augmentation techniques to improve the model's performance and demonstrate the effectiveness of CrackNet on the CFD2014 dataset.

In order to improve the amount and quality of a training dataset and create a more effective deep learning model, data augmentation is an essential deep learning technique. In order to overcome the scarcity of training data and eliminate overfitting, Islam et al. (2022) used data augmentation and transfer learning. The first suggested data augmentation strategy is the Random-Resized-Crop Method, which randomly selects a section of the image to crop and resizes it to the required size. The second method, known as the Random-Rotation Method, rotates the image across a variety of random angles. Angles between 15 degrees and +15 degrees are chosen for our training dataset. The Color-Jitter Method, which randomly modifies the brightness, is the third data augmentation technique employed.

4.4. AI-based classification algorithms in crack detection

AI-based crack classification strategies employ various techniques to identify and classify image cracks accurately. One approach is image patch crack classification, where the image is divided into smaller patches, and each patch is analyzed independently to determine if it contains a crack. Another strategy involves boundary box regression, which aims to detect the presence of cracks by predicting the bounding boxes that encompass them. This allows for precise localization and identification of crack objects within the image. Additionally, semantic segmentation is employed to classify each pixel in the image as either crack or non-crack, enabling a fine-grained understanding of crack distribution. By combining these strategies, AI systems can detect, classify, and analyze cracks, facilitating efficient and reliable inspection and maintenance in various domains such as infrastructure, manufacturing, and construction.

4.4.1. Image patch classification

In recent years, there has been a growing interest in data-driven approaches for various applications. This approach utilizes machine learning techniques to analyze data and make predictions or decisions based on the insights generated from the data. In this summary, we will discuss several papers that focus on data-driven approaches and their applications. Table 4 represents the different parameters that affect the accuracy of the AI-based crack detection system.

The paper by Hammouch et al. (2022) proposes an automated methodology for detecting and classifying cracks in Moroccan flexible pavements using Convolutional Neural Networks (CNN) and transfer learning. The periodic inspections of Moroccan roads using high-resolution cameras and GPS/DGPS receivers have been conducted since 2011, but the manual processing of pavement surface image sequences is complex, time-consuming, and subjective. The proposed approach shows good crack detection and classification results using both the CNN and pre-trained VGG-19 models. The study demonstrates the potential of data-driven approaches, specifically CNN and transfer learning, in automating crack detection and classification to diagnose road networks effectively.

The paper by Amhaz et al. (2016) proposes a new algorithm for automatic crack detection from 2D pavement images using minimal path localization. The approach selects a set of minimal paths and introduces two post-processing steps to improve detection quality. The algorithm considers pavement images' photometric and geometric characteristics and is validated on synthetic and natural images from different acquisition systems. The proposed method is compared to five existing methods and is found to provide very robust and precise results in a wide range of situations in a fully unsupervised manner, surpassing the current state of the art. The study highlights the potential of data-driven approaches in improving crack detection in pavement images. The survey of Padsumbiya et al. (2022) proposes a method for automatic crack detection on concrete surfaces using a simple Convolutional Neural Network (CNN) and compares it with a

TABLE 4 AI based classification techniques for crack detection.

References	Data size	Algorithm	Layer and epoch	Evaluation metric used	Classification result	Diversity in images
Silva and Lucena (2018)	3,500 samples	CNN	16 layers, 48 epoch, 3 parameters	Accuracy	Crack, No crack	Yes
Amhaz et al. (2016)	269 real images	Dijkstra	NA-layers, NA-epoch,5 parameters	DICE similarity coefficient	transverse cracks, longitudinal cracks, and alligator cracks	No
Padsumbiya et al. (2022)	40,000	CNN	4 layer, 5 epoch, 954,241 parameters	Precision, recall, and F1 score	Crack, No-crack	Yes
Ali et al. (2021)	25,000 samples	CNN	3 layers,20 epoch,2.7 million paramters	Accuracy, precision, recall, and F1-score	Crack, No-crack	Yes
Gopalakrishnan et al. (2017)	760 samples	VGG-16 DCNN	24 layers, 50 epochs, 144 million parameters	Accuracy, precision, recall, F1-score and Cohen's Kappa score	Crack, No-crack	No
Kim et al. (2021)	40,000 images	CNN	7 layers,20 epochs,5 parameters	Accuracy score, precision, recall, and F-Measure	Crack, No-crack	No
Maniat et al. (2021)	48,000 images	CNN	22 layers,100 epoch,20 million parameters	Precision, Recall, F1-Score	Transverse cracks, longitudinal cracks, and alligator cracks	No
Chen C. et al. (2021)	NA	SVM	NA	Accuracy	Transverse cracks, longitudinal cracks, and alligator cracks	Yes
Tabassum (2022)	370 Images	CNN	7 layers, 30 epoch, 813,475 parameters	Mean absolute error (MAE), mean square error (MSE), root mean squared error (RMSE), and r 2 value	Multiclass damage	No
Hammouch et al. (2022)	3,287 images	CNN VGG19	19 layers,40 epoch, 2,00,996,51 parameters	Precision, Recall, F1-Score	Longitudinal crack, transverse crack, alligator crack	Yes

Feed-Forward Fully Connected Neural Network. Using low-pixel density images, the proposed CNN is trained to detect and segregate cracked and non-cracked concrete surfaces. The model uses Max Pooling and optimization techniques and achieves a final accuracy of 97.8%. The study confirms the impact of Artificial Intelligence in Civil Engineering, where a simple neural network can carry out automatic crack detection, eliminating the need for high-cost digital image-capturing devices.

Ali et al. (2021) present a customized convolutional neural network (CNN) for crack detection in concrete structures and compare its performance with four existing deep learning methods based on training data size, data heterogeneity, network complexity, and number of epochs. The proposed CNN model and VGG-16 outperformed the other methods in terms of classification, localization, and computational time on a small amount of data. The evaluation considered various measures such as accuracy, precision, recall, and F1-score. The results indicate that increasing the training data size and reducing diversity reduced generalization performance and led to overfitting. The proposed CNN model and VGG-16 demonstrate superior crack detection and localization for concrete structures. The paper by Gopalakrishnan et al. (2017) suggests using a Deep Convolutional Neural Network (DCNN) trained on ImageNet to automatically detect cracks in Hot-Mix Asphalt (HMA) and Portland Cement Concrete (PCC) surfaced pavement images, which also contain non-crack anomalies and defects. The study aims to train a classifier on combined HMA-surfaced and PCC-surfaced photos that have different surface characteristics. A single-layer neural network classifier (with "Adam" optimizer) trained on ImageNet pre-trained VGG-16 DCNN features yielded the best performance.

Kim et al. (2021) propose a shallow convolutional neural network (CNN) for surface concrete crack detection that can be employed using low-power computational devices. The proposed LeNet-5 architecture is optimized and trained using 40,000 images in the METU dataset, achieving a maximum accuracy of 99.8% with minimum computation. The model's hyperparameters are optimized for crack detection, and its performance is compared with various pre-trained deep-learning models. The study concludes that the proposed method can be incorporated with autonomous devices, such as crewless aerial vehicles, for real-time inspection of surface cracks. In this paper, Prasanna et al. (2014) present a new automated crack detection algorithm called the STRUM classifier, which is used to detect cracks on concrete bridges. The algorithm uses machine learning classification and multiple visual features that are spatially tuned to potential crack regions, and it employs robust curve fitting to spatially localize likely crack regions even in the presence of noise. The algorithm is demonstrated using a state-of-the-art robotic bridge scanning

system and actual bridge data from hundreds of crack regions over two bridges. The results show that the STRUM classifier outperforms a more typical image-based approach, with a peak performance of 95% accuracy. The crack density map for the bridge mosaic, provides a computational description and a global view of the spatial patterns of bridge deck cracking.

The study of Maniat et al. (2021) investigates using Google Street View (GSV) pavement images to evaluate pavement quality. A convolutional neural network (CNN) is developed to perform image classification on GSV pavement images by dividing them into smaller image patches and classifying them into different categories of pavement cracks. The study compares the results of the CNN with those of a commercial visual inspection company and shows that GSV images can be effectively used for pavement evaluation. The designed CNN is able to classify pavement images into different crack categories. The paper by Amhaz et al. (2016) presents a new algorithm for automatic crack detection from 2D pavement images, which relies on localizing minimal paths within each image. The proposed approach selects a set of minimal tracks and introduces two post-processing steps to improve the detection quality. Compared to five existing methods, the algorithm is validated on synthetic and natural images from five different acquisition systems. The results show that the proposed algorithm provides very robust and precise results in a wide range of situations in a fully unsupervised manner, which is beyond the current state of the art.

The article by Silva and Lucena (2018) describes the development of a machine learning-based model to detect cracks on concrete surfaces using a deep learning convolutional neural network (CNN) image classification algorithm. The model is intended to increase the level of automation on concrete infrastructure inspection when combined with crewless aerial vehicles (UAV). A relatively heterogeneous dataset with 3500 images of concrete surfaces, balanced between images with and without cracks, were used for this work. The model's accuracy were recorded for different experiments, and the best investigation yielded a model with an accuracy of 92.27%, showing the potential of using deep learning for concrete crack detection.

4.4.2. Boundary box regression in crack detection

Object detection is another crucial aspect of AI-based crack detection systems. It involves identifying and localizing objects of interest in an image or video. In the case of crack detection, the thing of interest is the crack itself. Object detection is usually performed using deep learning models such as YOLO (You Only Look Once) and Faster R-CNN (Faster Region-based Convolutional Neural Networks). These models use convolutional neural networks (CNNs) to extract features from the input image and then use these features to identify and localize the crack. Table 5 represents the review articles published on crack detection with an object detection approach.

Fang et al. (2020) present a hybrid approach to detecting cracks in raw images using deep learning models and Bayesian probabilistic analysis. The technique involves retraining an object detector to identify crack patches with a suitable signal-to-noise ratio, generating ground truths of crack patches using a semi-automatic method, and using a Bayesian integration algorithm to suppress false detections. The algorithm uses a deep CNN to recognize the orientation of the crack segment in each detected patch, computes a Bayesian probability based on the accumulated evidence from seen adjacent patches within a neighborhood, and suppresses the patch lacking local supports as false detection. The proposed approach is evaluated on a comprehensive dataset of crack images and outperforms the state-of-the-art baseline approach on deep CNN classifier. Ablation experiments are conducted to demonstrate the effectiveness of the proposed techniques.

Hu et al. (2021) addressed the challenges in detecting cracks on asphalt pavement using traditional methods, which can be inefficient and miss detection. The study proposed using object detection based on the deep learning model YOLOv5 for pavement crack detection. The researchers collected 3,001 images of crack pavement with varying severity levels and used the YOLOv5 series models for training and testing. The results showed that the YOLOv5l model had the highest detection accuracy of 88.1%, and the YOLOv5s model had the shortest detection time of only 11.1 ms per image. The proposed approach could effectively detect cracks on asphalt pavement, which can improve road safety by identifying and repairing these cracks in time. Machine learning-based models can be difficult to generalize for various cracks, requiring the artificial design of pavement crack characteristics.

Li et al. (2019) propose a new method for crack inspection in aircraft structures, as existing methods are time-consuming and inaccurate. Their approach, called YOLOv3-Lite, utilizes depth-wise separable convolution, feature pyramids, and YOLOv3 to improve crack detection. By using depth-wise separable convolution, crack features are extracted, and parameters are reduced. The feature pyramid joins semantically strong features at high resolution for rich semantics, while YOLOv3 is used for bounding box regression. The results show that YOLOv3-Lite is 50% faster than YOLOv3 without any loss of detection accuracy, making it a state-of-the-art performance crack detection method for aircraft structures such as fuselage or engine blades.

Wan et al. (2021) introduce a solution to the issues of complex models and computational time-consuming problems in current deep learning-based methods for road damage detection. The proposed YOLO-LRDD model is a lightweight version of YOLOv5s that incorporates Shufe-ECANet with an ECA attention module as a new backbone network and utilizes BiFPN for reliable detection. The Focal-EIOU approach is applied in the training phase to obtain higher-quality anchor boxes, and the RDD2020 dataset is augmented with Chinese road scene samples for testing. Experimental results demonstrate that YOLO-LRDD outperforms several state-of-the-art object detection techniques in terms of accuracy and efficiency. Compared to YOLOv5s, YOLO-LRDD reduces model size by 28.8%, improves single image recognition speed by 22.3%, and is easier to implement in mobile devices due to its smaller and lighter model.

The paper by Hong et al. (2022) introduces the AugMoCrack network, a new method for detecting cracks at the level of bounding boxes. This approach focuses on identifying the location of crack objects through a morphological perspective, which sets it apart from other methods that require pixel-level

TABLE 5 Review of road crack object detection methods.

Author	Types of cracks	Algorithm	Additional features	Advantage	Limitations
Hu et al. (2021)	Fatigue crack, longitudinal crack, and transverse crack: object detection	YOLOv5 series	Backbone modules include Focus and CSP structures (neck modules include FPNs and PANs). The prediction module uses GIOU_Loss	Detection accuracy of >85%.	Model can be enhanced by combining semantic segmentation technology. Anchor-free object detection is not used
Li et al. (2019)	Aircraft cracks	YOLOv3-Lite	Combines depth wise separable convolution, feature pyramids, and YOLOv3	It can run on a mobile device due to its light-weight characteristics, YOLOv3-Lite is 50% faster than YOLOv3	Detection accuracy is less compared to YOLOv3
Wan et al. (2021)	Longitudinal cracks, lateral cracks, alligators, and potholes	YOLOv5	Shuffle-ECANet by adding an ECA attention, BiFPN Focal-EIOU in order to get higher-quality anchor box.	Sample imbalance treatment enhances the detection ability of the model for Small objects to a certain extent	Optimizing the sample imbalance processing method, can improve the sensitivity to small objects
Hong et al. (2022)	Cracks:, Patch, segmentaion	YOLOv5 Frequency Augmentation Morphological Filtering Morphological attention loss	A weakly supervised crack detection method with data augmentation. LossCA, which considers the connectivity of the detected crack, and LossPA, which helps in the strict box detection of the proposed method	Tightly detecting the bounding box area while reducing false duplicate detections. Excellent crack-detection performance even with little crack-training data	Not efficient for handling the road images with different textures
Ren et al. (2022)	Longitudinal crack, transverse crack, alligator crack, and pothole:	YOLOV5-CoordAtt	The main novelty of the proposed method is that various attention modules were introduced into the original networks to improve the ability of extracting the features of pavement cracks.	Precision of this method is 4.36% higher than the original model	The confidence scores of pavement cracks are not so high. Increase in computational load
Fang et al. (2020)	Images are night view, wet pavement, distance view, marble base,	Faster-RCNN	A DCNN is introduced to recognize the crack orientation. A two-phase Bayesian integration algorithm is proposed to verify the connected true detections and suppress false detections	To retain high sensitivity to small cracks, the dense outputs of overlapping patch detections from Faster R-CNN are employed.	Bayesian integration increases the computational cost
Zhao et al. (2022)	Longitudinal, transverse, and bifurcation cracks	Faster-RCNN	Combined with sparse representation and compressed sensing, the dataset was preprocessed, and various preprocessing algorithms were compared	Crack detection effects in complex situations such as road marking interference, shallow cracks multiple cracks, and ambiguity were significantly improved.	Type and severity of damage (cracking) is not considered

segmentation. The Poisson blending and high-frequency discrete cosine transform-based features to augment their training data. The network also employs morphological attention loss functions, considering neighbor connectivity and box area border to enhance detection accuracy. The researchers trained their network using two datasets and achieved a 4.5 and 2.5% increase in mean average precision (mAP) in the concrete crack and Crack500 datasets, respectively, compared to the baseline architecture. In a weakly supervised learning environment where training data is limited, the AugMoCrack network outperforms state-of-the-art crack detection methods.

Ren et al. (2022) present a pavement crack detection method that employs YOLOv5 as the base model for real-time inspection. Attention modules were added to improve the accuracy of deep learning-based methods since these methods typically need help to achieve high accuracy when dealing with small-sized pavement cracks. The proposed method used self-built datasets from Linyi city, and the results demonstrated that YOLOv5-CoordAtt with

attention modules had a precision of 95.27%, which was higher than other conventional and deep learning methods. The proposed method proved accurate in detecting pavement cracks in various situations. In another study, Zhao et al. (2022) offer a deep learning-based approach to seeing road cracks that utilize image sparse representation and compressed sensing to preprocess the datasets. This method achieves high accuracy and efficiency in crack identification and is robust when dealing with various road crack images. The method was evaluated using different algorithms, and the results showed that it outperforms the original method by increasing the mean average precision (mAP) by up to 5%.

4.4.3. Semantic segmentation in crack detection

Segmentation is dividing an image into multiple segments or regions based on predefined criteria such as color, texture, or shape. In the case of crack detection, segmentation is used to identify and highlight the crack region in an image. Several segmentation techniques are available, such as thresholding, edge detection, and deep learning-based segmentation methods, such as U-Net and Fully Convolutional Networks (FCNs). Table 6 illustrates the segmentation methods used in the literature with a combination of convolutional and deconvolution layers to generate a pixel-wise segmentation map, which can be used to quantify the severity of the crack.

Fan et al. (2021) propose a pothole detection approach using single-modal semantic segmentation, employing a convolutional neural network to extract visual features, a channel attention module to enhance consistency, and an atrous spatial pyramid pooling module to integrate spatial context information. They also use a multi-scale feature fusion module to reduce the semantic gap between different feature channel layers. Qu et al. (2021) present a network model for automatically detecting cracks with uneven strength from a complex background, which uses hierarchical feature fusion and connected attention architecture to recover lost details and incomplete extracted cracks. The model uses an improved DCA-SE-ResNet-50 as the backbone network and combines depthwise separable convolution and dilated convolution for crack feature fusion. Moreover, they design an attention layer to integrate feature map2 with feature map4 and incorporate feature maps of the low and high convolutional layers in the side network to assist in obtaining the final prediction map. The proposed method achieves state-of-the-art performances with the best F-score over 0.86 and 12 FPS on CFD, Crack500, and DCD datasets, and experimental results demonstrate the effectiveness of both methods.

Ai et al. (2018) proposed a new automatic pavement crack detection method that leverages multi-scale neighborhood information and pixel intensity using a probabilistic generative (PGM) based practice. The PGM calculates the probability of a crack for each pixel, and a fusion algorithm merges the probability maps from PGM and SVM approaches into a fused map that detects cracks more accurately than any of the original probability maps. Additionally, a weighted dilation operation is proposed to improve crack continuity. The proposed method outperforms state-of-the-art pavement crack detection algorithms regarding precision, recall, f1-score, and receiver operating characteristics. In contrast, Yang et al. (2019) address the challenge of automatic pavement crack detection by proposing a new network architecture called Feature Pyramid and Hierarchical Boosting Network (FPHBN) that integrates context information with low-level features and uses nested sample reweighting during training to balance the contributions of easy and hard samples to the loss. The paper also introduces an average intersection over union (AIU) measurement for crack detection evaluation.

Lin et al. (2023) proposed a deep learning based end-to-end segmentation approach focusing on the contextual semantic information and edge information on crack images. Three factors, including the extraction of multi-scale feature information, the spatiotemporal attention mechanism, and pyramidal pooling (PSA-Net), were used to construct this technique. Without expanding the number of network parameters, PSA-Net is a compact pavement crack detection model. By layered sample weighting to equalize the loss caused by simple and complex samples, it increases the accuracy of autonomous road crack detection.

4.5. Crack quantification

Crack quantification is the critical step after the crack skeleton is obtained from road image processing. Crack parameters such as length, width, and depth can be calculated to estimate the severity of the crack. Estimating crack severity will guide road maintenance officers to take the necessary actions to avoid mishaps. Figure 8 shows the crack severity categorization as low, medium, or high depending on the calculated values of the crack's length, width, and depth. These severity levels help in evaluating how the crack might affect the structural integrity of the civil infrastructure.

- Low severity: The absence or presence of relatively few interconnecting cracks with a width of 6 mm and a distance between cracks of $<$0.328 m. This type of crack can be so tight that determining its diameter might be difficult, if not impossible.
- Medium severity: The affected road section has a comprehensive, interconnected network of cracks. Crack widths range from 6 mm to 19 mm; they can also relate to any damage with an average width of $<$19 mm and near a pattern of mild severity cracking. The distance between cracks in this group is no more than 150 mm.
- High severity: The road area's crack pattern comprises moderately or severely developed linked cracks. Cracks bigger than 19 mm or any crack wider than 19 mm but near to medium to high-severity random cracking are associated with high severity.

Pavement cracks are a common problem on roads and highways, and their monitoring and quantification are essential for maintaining safe and efficient transportation networks. However, the complex texture of these cracks and the potential for noise and illumination to interfere with measurement accuracy have made traditional crack quantification methods challenging. The article by Sun et al. (2022) proposes a road crack monitoring and quantification method based on vehicle video to overcome

TABLE 6 Review of crack semantic segmentation[[Inline Image]] methods.

Author	Types of cracks	Algorithm	Additional features	Advantage	Limitations
Fan et al. (2021)	Pothole: semantic segmentation	Encoder Decoder	Unet with ResNet as Backbone Channel attention module Multi scale feature fusion model	The quantitative comparisons method achieves the state-of-the-art (SoTA) performance on both RGB images and transformed disparity images, outperforming three SoTA single-modal semantic segmentation networks	Cost is more Crack parameters not calculated
Qu et al. (2021)	Cracks: semantic segmentation	DCA-SE-ResNet-50	Feature fused (FF) module Attention layer Deeply-Supervised Nets (DSN)	Insensitive to noise crack marking, and can effectively distinguish stains and obstruction	Not very effective in detecting thin cracks
Ai et al. (2018)	Cracks-pixel level segmentation	A probabilistic generative model (PGM).	Support vector machine (SVM) multi-scale neighborhoods Information, Fusion Algorithm weighted Dilation to optimize the detected cracks	Maintains the continuity of crack	Produces fewer noise
Li et al. (2022b)	Transversal and longitudinal cracks, alligator Cracks: Segmentation	U-shaped network with the ResNet encoder	EDA: the high-efficiency dual attention module Multiscale attention hesitant fuzzy set (HFS) for crack severity classification	More suitable for the detection of crack images with unbalanced aspect ratios	Necessary to improve the attribute index and optimize the classification method of crack images, especially for classifying massive cracks and cracked cracks.
Yuan et al. (2022)	Cracks: Segmentation	Encoder Decoder	Residual detail attention (RDA) Attention gating strategy Cross-entropy loss function	Accuracy of crack extraction is improved with suppressed background artifacts	This method fails to achieve the desired results when dealing with fine cracks Integrity of fracture segmentation (e.g., blind inpainting), can be improved to extract more accurate and complete features

FIGURE 8
Crack severity levels.

the limitations of traditional methods. The method includes automated vehicle-mounted equipment with GPS signals to capture crack images with location information, extracting morphological features of dynamic road cracks, and a calculation algorithm based on the United Kingdom scanning grid and projection method. The proposed method also improves the crack distress evaluation method through the analysis of different crack grades. The experimental results indicate strong reliability and adaptability with high-frequency and wide-range road detection, and the proposed method has the potential to improve the monitoring and quantification of pavement cracks and ensure the safety and longevity of transportation infrastructure.

On the other hand, Matarneh et al. (2023) developed an automated tool using the Hough transform algorithm for detecting and classifying pavement cracks to optimize road maintenance and prevent possible failures. The article reviews existing attempts to use the algorithm and proposes a simple, low-cost method that achieves high accuracy for detecting and classifying vertical, diagonal, and horizontal cracks. The article suggests that this low-cost image processing method has the

potential to automate pavement crack detection and guide long-term pavement maintenance decisions, which can reduce costs for highway agencies. The originality of the article lies in its successful testing of the Hough transform algorithm for automated cracks and distresses classification.

The article by Avendaño (2020) discusses the challenges of manual inspections for assessing damages in civil engineering structures. It highlights using image-based checks using cameras or crewless aerial vehicles (UAV) combined with image processing to overcome these challenges. The article presents an approach combining different aspects of the inspection, from data acquisition through crack detection to quantifying essential parameters. A convolutional neural network (CNN) is used to identify cracks, and different quantification methods are explored to determine the width and length of the damages. The results demonstrate a low to no false negative rate for crack identification using the CNN and the highest accuracy estimation for 0.2 mm cracks during quantification.

Similarly, Deng et al. (2023) propose an integrated framework for automatically detecting, segmenting and measuring road surface cracks. The approach involves using the YOLOv5 algorithm for crack detection and a modified Res-UNet algorithm for accurate segmentation at the pixel level—a novel crack surface feature quantification algorithm to determine the width and length of the cracks. The proposed method is validated using a road crack dataset containing complex environmental noise, and it shows higher accuracy for crack segmentation under complex backgrounds compared to other methods. The developed crack surface feature algorithm has an accuracy of 95% in identifying the crack length and a root mean square error of 2.1 pixels in identifying the crack width, with the accuracy being 3% higher in length measurement than that of the traditional method.

Ha et al. (2022) propose an integrated framework for the automated detection, classification, and severity assessment of road cracks to optimize pavement management systems. The proposed system expands the number of detected crack types to five (alligator, longitudinal, transverse, pothole, and patching). It includes the assessment of crack severity, which typically needs to be improved in related studies. The studied research article uses SqueezeNet, U-Net, and Mobilenet-SSD models to achieve an accuracy of 91.2% for both crack type and severity assessment. The proposed system uses U-Nets for linear and area cracking to improve object detection performance and automate the evaluation of crack severity. The suggested automated pavement management system better reflects each country's requirements for various crack types and severity standards.

Moreover, Carrasco et al. (2021) suggest a novel automated method for measuring the width of surface cracks in civil engineering infrastructure. The traditional visual inspection method and manual measurement with a crack-width comparator gauge are time-consuming and error-prone. Although algorithms for automatic crack detection have been developed, most still need to address the problem of crack width evaluation. The proposed method consists of three stages: anisotropic smoothing, segmentation, and stabilized central points by k-means adjustment. It allows the characterization of both crack width and curvature-related orientation and has been validated by assessing the surface cracking of fiber-reinforced earthen construction materials. The preliminary results show that the proposal is robust, efficient, and highly accurate at estimating crack width in digital images, effectively detecting natural cracks as small as 0.15 mm width regardless of the lighting condition.

4.5.1. Method for calculating the road segment severity index

Boucetta et al. (2021) suggested three indices combined in the severity index computation unit: the alligator cracks index (ACI), the transverse cracks index (TCI), and the longitudinal cracks index (LCI). These indices are calculated based on the cracks observed in the gathered photos. The road network creation stage entails creating a graph from the collected road network data. The weighted network creation step generates a weighted graph based on the collected severity indices and road data. The severity indices are calculated using the edge weights of the road network

The road segment severity index (SI), is a composite of three indices: the ACI (Eq. 2), TCI (Eq. 3), and LCI (Eq. 4) of the same segment. Each of the three indices takes into account the number of cracks discovered in a road segment and their severity level, which is enabled by the suggested detection and classification methods discussed in the preceding divisions.

$$SI\,(i) = \frac{ACI(i) + TCI(i) + LCI(i)}{3} \tag{1}$$

$$ACI\,(i) = \frac{Cf_1 \times LAC + Cf_2 \times MAC + Cf_3 \times HAC}{\sum_{j=1}^{3} Cf_j} \tag{2}$$

LAC, MAC, and HAC denote the number of low, medium, and high severity alligator cracks in road segments, respectively.

$$TCI\,(i) = \frac{Cf_1 \times LTC + Cf_2 \times MTC + Cf_3 \times HTC}{\sum_{j=1}^{3} Cf_j} \tag{3}$$

LTC, MTC, and HTC denote the number of low, medium, and high severity transverse cracks in road segments, respectively.

$$LCI\,(i) = \frac{Cf_1 \times LLC + Cf_2 \times MLC + Cf_3 \times HLC}{\sum_{j=1}^{3} Cf_j} \tag{4}$$

LLC, MLC, and HLC are the counts of the low, medium, and high severity longitudinal cracks in road segments, respectively, and are coefficients associated with each cost parameter in each equation independently. Experts in the area are in charge of verifying the values of these coefficients to improve the optimization of the weight calculation algorithm.

5. Learning algorithms for crack detection

Learning algorithms are an essential part of AI-based crack detection systems. The primary goal of learning algorithms is to provide the system with the ability to learn from a training dataset to recognize and classify the crack images. There are three

types of learning algorithms: supervised, unsupervised, and semi-supervised. In supervised learning, the system is trained using labeled data, whereas, in unsupervised learning, the system is trained without labeled data. In semi-supervised learning, the system is trained with labeled and unlabeled data. The choice of learning algorithm depends on the availability of labeled data and the type of problem to be solved.

5.1. Supervised learning

Qu et al. (2022) introduce a new method for detecting cracks in concrete structures using deep learning and multiscale fusion techniques. The use of modified U-Net architecture achieves high accuracy rates of up to 98.67% on a diverse dataset of concrete images with various types of cracks. The study emphasizes the potential of combining deep learning and multiscale fusion techniques for accurate and efficient crack detection in concrete structures. Zou et al. (2018) proposed an end-to-end deep learning-based approach for crack segmentation in pavement images.VGG-16 and U-Net networks extract and refine hierarchical features for precise crack segmentation. The proposed method outperforms traditional methods and achieves state-of-the-art results on benchmark datasets. The study highlights the potential of deep learning-based approaches for accurate and efficient crack segmentation in pavement images. Zhang et al. (2016) present an automatic crack detection method for concrete images using a convolutional neural network (CNN). CNN architecture with two convolutional and two fully connected layers achieves a high accuracy rate of 94.4% on a dataset of concrete images with cracks. The study demonstrates the potential of deep learning-based approaches for automatic crack detection in concrete images. Chen Y. et al. (2021) and Sun et al. (2021) comprehensively review image processing-based techniques for crack detection in concrete pavements. The studied article examines crack detection techniques, including thresholding, edge detection, texture analysis, and machine learning-based methods. The study emphasizes the importance of developing efficient and accurate techniques for crack detection in concrete pavements to ensure their safety and durability. It provides insights for future research in the field.

5.2. Unsupervised learning

Unsupervised learning is a machine learning technique involving training models on data without explicit supervision or labeled data. Several studies have explored unsupervised learning techniques for various applications. Duan et al. (2020) proposed an unsupervised deep-learning framework for anomaly detection in industrial processes. The use of generative adversarial network (GAN) to learn the normal data distribution and identify anomalous samples. The proposed method outperformed traditional anomaly detection techniques and achieved high accuracy rates. The study demonstrates the potential of unsupervised learning techniques for anomaly detection in industrial processes.

Similarly, Li et al. (2021) proposed an unsupervised deep learning-based approach for anomaly detection in power systems. The variational autoencoder (VAE) learn the expected behavior of the power system and identify anomalies. The proposed method outperformed traditional anomaly detection techniques and achieved high accuracy rates. The study highlights the potential of unsupervised learning techniques for anomaly detection in power systems. Wu et al. (2021) proposed an unsupervised deep-learning framework for image classification. An autoencoder-based clustering approach group images with similar features and achieved high accuracy rates on benchmark datasets. The proposed method outperformed traditional clustering techniques and demonstrated the potential of unsupervised learning techniques for image classification.

Mubashshira et al. (2020) proposed an unsupervised learning-based approach for feature extraction in human action recognition. Use of a stacked denoising autoencoder (SDAE) to learn discriminative features from raw sensor data and achieved high accuracy rates on benchmark datasets. The proposed method outperformed traditional feature extraction techniques and demonstrated the potential of unsupervised learning techniques for human action recognition.

Methods for supervised crack detection today mainly rely on labeled data. Li et al. (2022b) presented an unsupervised reconstruction-based concrete crack detection approach based on nnU-Net. This approach works better when the normal sample distribution's variance is modest. In this instance, while the abnormal image cannot be entirely learned from the semantic information of the normal sample, the normal image can be improved rebuilt. In order to rebuild the original image, the input image first goes to the trained model, and the output image is then utilized as the input image of the final model. Following study, he suggested a reconstruction approach, and his selection of the loss function encourages the network to better rebuild the system.

Li et al. (2021) study's main goal is to address the weak generalization and low intelligence issues with crack detection. Considering efficiency and model simplicity, the studied article proposes a fused deep neural network model architecture with the K_means clustering algorithm. The K_means clustering technique is utilized to generate the pseudolabels that the AlexNet model uses to train its model, and this architecture is based on the original AlexNet model. This fused architecture's key benefit is that it avoids the disadvantages of supervised learning techniques by not requiring manually labeled ground truth images for model training. This model performs satisfactorily after being trained on cracked images acquired from cellphones and automated cars in a variety of environmental scenarios with different image quality.

5.3. Semi-Supervised learning

Semi-supervised learning is a machine learning approach that uses labeled and unlabeled data to improve model accuracy. In pavement evaluation and maintenance planning, detecting cracks is a critical task. The paper by Tang et al. (2022) discusses using Artificial Neural Networks (ANNs) for crack detection in pavement evaluation and maintenance planning. Most existing models use

the Fully Supervised Learning (FSL) approach, which relies on high-quality annotation for reasonable accuracy. However, this approach is costly and time-consuming, especially for complex networks. This paper proposes a Weakly Supervised Learning U-Net (WSL U-Net) for pavement crack segmentation. This approach uses weakly labeled images to train the network, significantly reducing the labor cost and human involvement in image annotation. The experimental results show that WSL U-Net outperforms some Semi-Supervised Learning (Semi-SL) and WSL methods and achieves comparable performance with its FSL version. The dataset cross-validation also demonstrates that WSL U-Net is more robust with fewer overfitting concerns and better generalization capability.

Detecting cracks at the pixel level is vital for building and road inspections, but it requires time-consuming pixel-level annotations. Previous work proposed a weakly-supervised approach but struggled with lighter-colored cracks and non-crack targets. Inoue and Nagayoshi (2023) proposed a data-driven annotation refinement approach, which is effective regardless of a dataset's pixel brightness profile. The method speeds up the annotation process by factors of 10 to 30 while maintaining detection accuracy on three crack segmentation datasets and one blood vessel segmentation dataset.

Zhu and Song (2020) present a weakly supervised approach for detecting and segmenting cracks in asphalt concrete bridge decks, which are challenging to detect using conventional methods due to their dark color and complex nature. The proposed method uses an autoencoder to differentiate data and highlight unlabeled data features for weakly supervised learning. K-means clustering is used to classify features, and semantic segmentation is performed under weak supervision to identify cracks. The proposed method is evaluated on a dataset of six types of defects on asphalt concrete bridge decks and outperforms existing methods reported in the references.

TABLE 7 Challenges and future scope in crack detection.

Problem area	Challenges in crack detection	Future direction
Dataset	Imbalance image data samples	Transfer learning and semi-supervised GAN
	Heterogeneity in data samples	Extreme Learning Machine
	3D dataset	3D-Point-Cloud-Based Deep Neural Network
	Lacking labeled/annotated datasets	Deep active learning
Accuracy	Accuracy is compromised in varied road conditions like shadow, stains, oil spots etc.	Attention Modules Multi scale feature Fusion Combine object detection with Segmentation
Computational cost	Labeling cost Models are very large and heavy. It takes a lot of time and computing power to run Number of possible combinations for bounding boxes is huge, and these networks tend to be computationally demanding.	Anchor Free Object Detection Unsupervised deep learning models can be used
Real time system	Improve the performance in real time crack localization Need huge amounts of training data and powerful computational infrastructure	Distributed Lightweight Deep Learning Models
Crack quantification	Calculating the crack width, length, and depth Crack categorization based on severity	

6. Challenges and future scope

In this article, we reviewed various road imaging methods focusing on the data that transforms through different image processing phases and outputs the segmented crack with its severity level. With a data-driven approach, we studied pre- and post-processing techniques used to improve image quality. We also looked at learning techniques used for based crack classifications. The following problem areas related to crack detection are majorly discussed in Table 7 with their future scope.

Furthermore, future dataset collection efforts will prioritize accuracy in computer vision tasks. Attention modules will enable models to focus on relevant features, while multiscale feature fusion will capture information at different levels of detail. Dataset creation will also facilitate the integration of object detection with segmentation. By emphasizing these aspects, dataset collection efforts can drive advancements and enhance the accuracy of computer vision applications in crack detection.

Moreover, addressing computational costs in object detection algorithms will be crucial in future dataset collection. This may involve exploring anchor-free techniques as alternatives to reduce computational complexity. Using unsupervised deep learning models can also help mitigate annotation overhead and computational requirements. By focusing on these objectives, dataset collection efforts can lead to developing more efficient object detection models without compromising accuracy and performance in crack detection applications.

Integrating distributed lightweight deep learning models will play a significant role in the future of real-time systems for crack detection. These models, designed for resource-constrained devices, will enhance the efficiency and responsiveness of real-time applications, particularly in road pavement. By leveraging distributed computing and lightweight architectures, real-time systems can achieve improved performance, low-latency processing, and benefit domains such as edge computing and IoT devices.

Lastly, advancements in image analysis techniques, automated detection systems, and the integration of non-destructive evaluation methods will shape the future of crack quantification. The aim is to enhance accuracy, automate the quantification process, and enable real-time monitoring for proactive maintenance and improved infrastructure safety in road pavement.

6.1. Future scope

In the context of crack detection in road pavement, the future scope of dataset collection is expected to drive advancements in research and application. Specialized datasets will support emerging techniques such as transfer learning, semi-supervised GANs, extreme learning machines, 3D-point-cloud-based deep neural networks, and deep active learning. These datasets will be diverse and domain-specific, covering many complexities. The aim is to leverage pre-trained models, explore advanced learning algorithms, and advance the field of 3D deep learning while reducing the effort required for labeling. This focus on dataset creation holds the potential to foster innovation and progress in machine learning and artificial intelligence, specifically for crack detection in road pavement.

6.1.1. Deep active learning

Deep active learning for crack detection combines deep learning models with active learning strategies to improve the efficiency and effectiveness of crack detection. It involves iteratively selecting the most informative samples from an unlabeled pool, annotating the samples by experts, and incorporating the labeled data into the training set to update the model. This process reduces annotation effort, improves model performance, and lowers labeling costs. Deep active learning provides flexibility, adaptability, and opportunities for exploring different active learning strategies to enhance crack detection. Overall, it enables efficient and accurate crack detection by leveraging deep and active learning strengths. Lv et al. (2020) developed an active learning architecture to minimize the labeling work required for flaw identification. An iteration pattern is used to train the detection model in the proposed system. An uncertainty sampling approach chose input photos based on their uncertainty levels for annotation of image data. Also, they developed an average margin approach to determine the sample ratios among defect categories to verify the sampling number for annotations.

6.1.2. Generative adversarial networks

GANs can be used for crack detection by generating realistic synthetic crack images and training a discriminator to differentiate between genuine and artificial cracks. The adversarial training process improves the accuracy of crack detection models. GANs can augment training data, enable unsupervised learning, assist in domain adaptation, and support semi-supervised learning. By leveraging the generative and discriminative capabilities of GANs, crack detection models can be enhanced for more effective infrastructure analysis and maintenance. Shim et al. (2020) proposed an automatic method for detecting cracks in concrete structures that combines transfer learning and data augmentation techniques. Their approach achieved high accuracy rates of 96.3 and 94.3%, respectively and is more efficient than traditional methods. The study highlights the significance of data augmentation and transfer learning in enhancing the accuracy and robustness of their proposed approach. The suggested technique can be valuable for engineers and practitioners working in crack detection in concrete structures.

6.1.3. Meta-learning

Meta-learning, known as "learning to learn," can enhance crack detection systems by enabling models to adapt to new crack types rapidly, learn from limited labeled data, and transfer knowledge across different structures or materials. It facilitates the efficient adaptation of crack detection models to changing crack patterns, optimizing hyperparameters, and guiding active learning and sample selection. With meta-learning, models can quickly generalize from previous crack detection tasks, allowing for effective crack identification and classification with minimal training data. The ability to learn empowers crack detection systems to leverage prior knowledge and experiences, accelerating their performance in novel scenarios. While the application of meta-learning to crack detection is still in its early stages, these principles hold promise for improving crack detection systems' efficiency, adaptability, and accuracy, contributing to safer and more reliable infrastructure management. Mundt et al. (2019) research contributes to the field of concrete defect recognition by introducing new meta-learning approaches and highlighting their effectiveness in finding optimized CNN architectures. These architectures demonstrate improved accuracy and parameter efficiency performance, addressing the challenges posed by the complex and varied nature of concrete defects in real-world scenarios. Further research and exploration in this area can unlock the full potential of meta-learning in crack detection.

6.1.4. Attention module

In recent studies, attention mechanisms have been utilized to improve the accuracy of automatic pavement crack detection using deep learning models. Xiang et al. (2020) proposed an end-to-end trainable deep convolution neural network that incorporates attention mechanisms to detect pavement cracks accurately. The network architecture includes a pyramid and spatial-channel combinational attention modules to refine crack features. At the same time, dilated convolution is used to avoid losing crack details during pooling operations. In addition, the Lovász hinge loss function is used to train the model on the CRACK500 dataset and evaluate it on three pavement crack datasets. The results show that the proposed method outperforms other methods regarding experimental precision.

Jing et al. (2022) proposed the AR-UNet network model to improve further the accuracy of crack detection, which introduces a convolutional block attention module (CBAM) in the encoder and decoder of the U-Net. The CBAM allows for effectively extracting global and local detail information, while the basic block prevents network degradation and layer growth. The method is tested on multiple datasets and achieves higher crack detection accuracy than existing methods.

Furthermore, Yu et al. (2022) proposed a U-shaped encoder-decoder network, RUC-Net, for automatic pavement crack detection. The scSE attention module and focal loss function were incorporated into the network to enhance detection accuracy. The proposed method was evaluated on three public datasets, demonstrating superior performance to other methods such as FCN, Unet, and SegNet. Additionally, studies were conducted on the CFD dataset to compare the effectiveness of different scSE

modules and their combinations in improving the performance of crack detection.

Wan et al. (2021) proposed CrackResAttentionNet, an encoder-decoder network-based architecture with position and channel attention modules to accurately detect pavement cracks with complex textures and different lighting conditions. The architecture outperformed popular models such as ENet, ExFuse, FCN, LinkNet, SegNet, and UNet regarding precision, mean IoU, recall, and F1 for both public and self-developed datasets. Ren et al. (2022) proposed an automatic pavement crack detection method using YOLOV5 as the base model and employing attention modules to improve detection accuracy for small cracks. The proposed CoordAtt module selectively attends to relevant features for better crack detection. The proposed method was evaluated on self-built datasets and outperformed conventional and deep learning methods, with a precision of 95.27%. Adding attention modules can effectively enhance the ability of crack detection under various situations.

7. Discussion

Detecting cracks in civil structures is a critical task to ensure the safety and longevity of our infrastructure. The development of advanced image processing technologies has made it possible to detect and analyze cracks non-invasively. In this study, we reviewed various image processing technologies and investigated the crack detection method based on image processing.

One of the main challenges in crack detection is data collection. Sensor systems such as UAVs, camera-mounted vehicles, and smartphones can be used to collect high-quality data. However, the 3D dataset is not widely available, so we are limited to using 2D data for our research. Furthermore, road image data can have a lot of noise, including uneven illumination, road lanes, stains, etc. Therefore, it requires preprocessing techniques such as grayscale conversion, histogram equalization, filtering, morphological operations, etc. Although this process can be time-consuming, obtaining high-quality data for effective crack detection is necessary. Labeling the data is another crucial aspect of crack detection. It can be done at image, block, or pixel levels, and it can be a tedious task. Moreover, combining textures such as asphalt and cement in datasets is often tricky, and there is a lack of small datasets available for Indian roads. Therefore, we plan to generate a combined textures dataset that includes Indian roads. Kanaeva and Ivanova (2021) diversity of road image data comprises images of various road surfaces, such as asphalt, cement, etc. Models can more easily generalize to situations in the real world when synthetic datasets are quickly modified to include a diverse variety of road picture data. In order to generate a robust and diversified dataset containing damaged and undamaged instances, synthetic dataset generation is a technique that is frequently used in machine learning and computer vision. Researchers can produce synthetic images that reflect the traits of the damage they seek to identify by employing a variety of techniques, including data augmentation, generative models, or simulations.

Learning algorithms play a significant role in crack detection and supervised learning algorithms such as CNN VGG16, Google net, transfer learning, and semi-supervised learning algorithms that require partial labeling can be used. Unsupervised learning algorithms were rarely used in this field. Our research uses a semi-supervised learning approach to train our models effectively. Crack data augmentation is necessary to deal with the limited availability of small datasets. Generative adversarial networks (GANs) can be used to generate synthetic data for training, improving the models' accuracy.

Different crack classification approaches can be used at different levels, such as object detection with YOLO, Faster RCNN, segmentation thresholding, edge detection, and Unet FCN DEEPVLAB. However, real-time crack detection is still challenging, and accuracy is a significant problem due to the computational cost involved. To overcome this challenge, we plan to use object detection followed by segmentation, which has proven more effective.

Crack quantification is another essential aspect of crack detection. Crack parameters such as length, width, and depth need to be evaluated to determine the severity of the crack. However, depth calculation is complex, and there is a need for further research in this area. Finally, performance evaluation metrics such as F1 score, mean average precision (MAP), recall, precision, IOU, AUC, etc., are used to assess the accuracy of the models. We plan to use MAP and AUC as our evaluation metrics.

Hence, image processing-based crack detection is a promising technique for detecting and analyzing cracks in civil structures. Our research will address the challenges related to data collection, labeling, learning algorithms, crack classification, crack quantification, and performance evaluation metrics. Our findings will contribute to developing more accurate and efficient crack detection methods in the future.

8. Conclusions

This article comprehensively reviews AI-based image-processing technologies for crack detection in civil structures. The study explores and investigates the challenges in data-driven aspects of crack detection methods. The researchers carefully selected relevant research articles on crack detection systems and analyzed them in this review. The review first examines the element of data collection and dataset analysis. Researchers were observed to use camera-type images for research predominantly, and authentic datasets were predominantly utilized. This ensured efficiency and ease of implementation. The accuracy and error levels of the analyses were thoroughly assessed. The review also highlights methods relevant for future research in image processing-based crack detection systems. Machine learning (ML) and deep learning (DL) techniques have become mainstream technologies in developing more advanced pavement crack detection algorithms. DL-based approaches such as image patch classification, crack semantic segmentation, and boundary box regression models were compared using uniform evaluation metrics. Crack segmentation is identified as a significant research area in engineering, particularly in the field of image recognition technology. It plays a crucial role in extending the service life of civil structures and reducing safety hazards. While convolutional neural networks (CNNs) have achieved remarkable results in the segmentation and detection of road cracks, further improvements

are still necessary. The article extracts challenges in crack detection and suggests future directions for research. One area of focus is exploring methods to compress the network scale without compromising segmentation accuracy. Additionally, the physical quantification of structural cracks is an important aspect for further investigation.

Author contributions

DV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing—review and editing. PC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing—original draft, Writing—review and editing. SP: Data curation, Formal analysis, Investigation, Resources, Supervision, Validation, Visualization, Writing—review and editing. SM: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Validation, Writing—review and editing. KK: Formal analysis, Funding acquisition, Investigation, Resources, Supervision, Validation, Visualization, Writing—review and editing.

Funding

Funding is organized by the Symbiosis International University (SIU) Pune.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ai, D., Jiang, G., Kei, L. S., and Li, C. (2018). Automatic pixel-level pavement crack detection using information of multi-scale neighborhoods. *IEEE Access* 6, 24452–24463. doi: 10.1109/ACCESS.2018.2829347

Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., Serhani, M. A., et al. (2021). Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors* 21, 1688. doi: 10.3390/s21051688

Amhaz, R., Chambon, S., Idier, J., and Baltazart, V. (2016). Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection. *IEEE Trans. Int. Trans. Syst.* 17, 2718–2729. doi: 10.1109/TITS.2015.2477675

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., and Sekimoto, Y. (2021). RDD2020: an annotated image dataset for automatic road damage detection using deep learning. *Data Brief* 36, 107133. doi: 10.1016/j.dib.2021.107133

Avendaño, J. (2020). *Identification and Quantification of Concrete Cracks Using Image Analysis and Machine Learning*. KTH VETENSKAP OCH KONST.

Bayer, M., Kaufhold, M. A., and Reuter, C. (2023). A survey on data augmentation for text classification. *ACM Comput. Surv.* 55, 1–39. doi: 10.1145/3544558

Boucetta, Z., El Fazziki, A., and El Adnani, M. (2021). A deep-learning-based road deterioration notification and road condition monitoring framework. *Int. J. Int. Eng. Syst.* 14, 503–515. doi: 10.22266/ijies2021.0630.42

Byrne, J., O'Keeffe, E., Lennon, D., and Laefer, D. F. (2017). 3D reconstructions using unstabilized video footage from an unmanned aerial vehicle. *J. Imag.* 3, 15. doi: 10.3390/jimaging3020015

Cao, M. T., Tran, Q. V., Nguyen, N. M., and Chang, K. T. (2020). Survey on performance of deep learning models for detecting road damages using multiple dashcam image resources. *Adv. Eng. Inf.* 46, 101182. doi: 10.1016/j.aei.2020.101182

Cao, W., Liu, Q., and He, Z. (2020). Review of pavement defect detection methods. *IEEE Access* 8, 14531–14544. doi: 10.1109/ACCESS.2020.2966881

Carrasco, M., Araya-Letelier, G., Velázquez, R., and Visconti, P. (2021). Image-based automated width measurement of surface cracking. *Sensors* 21, 7534. doi: 10.3390/s21227534

Chen, C., Seo, H., Jun, C., and Zhao, Y. (2022). A potential crack region method to detect crack using image processing of multiple thresholding. *Signal Image Vid. Proc.* 16, 1673–1681. doi: 10.1007/s11760-021-02123-w

Chen, C., Seo, H., Jun, C. H., and Zhao, Y. (2021). Pavement crack detection and classification based on fusion feature of LBP and PCA with SVM. *Int. J. Pave. Eng.* 23, 3274–3283. doi: 10.1080/10298436.2021.1888092

Chen, Y., Ding, Y., Zhao, F., Zhang, E., Wu, Z., Shao, L., et al. (2021). Surface defect detection methods for industrial products: a review. *Appl. Sci.* 11, 7657. doi: 10.3390/app11167657

Chen, Y., Naud, C. M., Rangwala, I., Landry, C. C., and Miller, J. R. (2014). Comparison of the sensitivity of surface downward longwave radiation to changes in water vapor at two high elevation sites. *Environ. Res. Lett.* 9, 114015. doi: 10.1088/1748-9326/9/11/114015

Deng, L., Zhang, A., Guo, J., and Liu, Y. (2023). An integrated method for road crack segmentation and surface feature quantification under complex backgrounds. *Remote Sens.* 15, 1530. doi: 10.3390/rs15061530

Duan, L., Geng, H., Pang, J., and Zeng, J. (2020). "Unsupervised pixel-level crack detection based on generative adversarial network," in *Proceedings of the 2020 5th International Conference on Multimedia Systems and Signal Processing (ICMSSP '20)* (New York, NY: Association for Computing Machinery), 6–10.

Ersoz, A. B., Pekcan, O., and Teke, T. (2017). Crack identification for rigid pavements using unmanned aerial vehicles. *IOP Conf. Ser.: Mater. Sci. Eng.* 236, 012101. doi: 10.1088/1757-899X/236/1/012101

Fan, J., Bocus, M. J., Hosking, B., Wu, R., Liu, Y., Vityazev, S., et al. (2021). "Multi-scale feature fusion: Learning better semantic segmentation for road pothole detection," in *2021 IEEE International Conference on Autonomous Systems (ICAS)*. Piscataway, NJ: IEEE, 1–5.

Fang, F., Li, L., Gu, Y., Zhu, H., and Lim, J. H. (2020). A novel hybrid approach for crack detection. *Patt. Recognit.* 107, 107474. doi: 10.1016/j.patcog.2020.107474

Feng, X., Xiao, L., Li, W., Pei, L., Sun, Z., Ma, Z., et al. (2020). Pavement crack detection and segmentation method based on improved deep learning fusion model. *Mathematic. Prob. Eng.* 2020, 1–22. doi: 10.1155/2020/8515213

Fu, R., Xu, H., Wang, Z., Shen, L., Cao, M., Liu, T., et al. (2020). Enhanced intelligent identification of concrete cracks using multi-layered image preprocessing-aided convolutional neural networks. *Sensors* 20, 2021. doi: 10.3390/s20072021

Ghaderzadeh, M., Aria, M., and Asadi, F. (2021). X-ray equipped with artificial intelligence: changing the COVID-19 diagnostic paradigm during the pandemic. *BioMed Res. International*, 2021 doi: 10.1155/2021/9942873

Ghaderzadeh, M., Aria, M., Hosseini, A., Asadi, F., Bashash, D., Abolghasemi, H., et al. (2022). A fast and efficient CNN model for B-ALL diagnosis and its subtypes classification using peripheral blood smear images. *Int. J. Int. Syst.* 37, 5113–5133. doi: 10.1002/int.22753

Gheisari, M., Ebrahimzadeh, F., Rahimi, M., Moazzamigodarzi, M., Liu, Y., Dutta Pramanik, P. K., et al. (2023). Deep learning: applications, architectures, models, tools, and frameworks: a comprehensive survey. *CAAI Trans. Int. Technol.* 8, 1–26. doi: 10.1049/cit2.12180

Golding, V. P., Gharineiat, Z., Munawar, H. S., and Ullah, F. (2022). Crack detection in concrete structures using deep learning. *Sustainability* 14, 8117. doi: 10.3390/su14138117

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., and Agrawal, A. (2017). Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construct. Build. Mat.* 157, 322–330. doi: 10.1016/j.conbuildmat.2017.09.110

Ha, J., Kim, D., and Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. *J. Supercomput.* 78, 17721–17735. doi: 10.1007/s11227-022-04560-x

Hamishebahar, Y., Guan, H., So, S., and Jo, J. (2022). A comprehensive review of deep learning-based crack detection approaches. *Appl. Sci.* 12, 1374. doi: 10.3390/app12031374

Hammouch, W., Chouiekh, C., Khaissidi, G., and Mrabti, M. (2022). Crack detection and classification in moroccan pavement using convolutional neural network. *Infrastructures* 7, 152. doi: 10.3390/infrastructures7110152

Hassan, S., O'Sullivan, D., Mckeever, S., Power, D., Mcgowan, R., and Feighan, K. (2022). "Detecting patches on road pavement images acquired with 3D laser sensors using object detection and deep learning," in *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) – Volume 5: VISAPP*, 413–420.

Hong, Y., Lee, S., and Yoo, S. B. (2022). AugMoCrack: augmented morphological attention network for weakly supervised crack detection. *Electr. Lett.* 58, 651–653. doi: 10.1049/ell2.12562

Hosseini, A., Eshraghi, M. A., Taami, T., Sadeghsalehi, H., Hoseinzadeh, Z., Ghaderzadeh, M., et al. (2023). A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: a design and implementation study. *Inf. Med. Unlocked* 39, 101244. doi: 10.1016/j.imu.2023.101244

Hsieh, Y. A., and Tsai, Y. J. (2020). Machine learning for crack detection: review and model performance comparison. *J. Comput. Civil Eng.* 34, 04020038. doi: 10.1061/(ASCE)CP.1943-5487.0000918

Hu, G. X., Hu, B. L., Yang, Z., Huang, L., and Li, P. (2021). Pavement crack detection method based on deep learning models. *Wireless Commun. Mob. Comput.* 2021, 1–13. doi: 10.1155/2021/5573590

Inoue, Y., and Nagayoshi, H. (2023). Weakly-supervised crack detection. *IEEE Trans. Int. Trans. Syst.* 24, 1–13. doi: 10.1109/TITS.2023.3286934

Islam, M. M., Hossain, M. B., Akhtar, M. N., Moni, M. A., and Hasan, K. F. (2022). CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack. *Algorithms* 15, 287. doi: 10.3390/a15080287

Jeong, J., Jo, H., and Ditzler, G. (2020). Convolutional neural networks for pavement roughness assessment using calibration-free vehicle dynamics. *Comput. Aided Civil Infrastr. Eng.* 35, 1209–1229. doi: 10.1111/mice.12546

Jing, P., Yu, H., Hua, Z., Xie, S., and Song, C. (2022). Road crack detection using deep neural network based on attention mechanism and residual structure. *IEEE Access* 11, 919–929. doi: 10.1109/ACCESS.2022.3233072

Jordan, S., Moore, J., Hovet, S., Box, J., Perry, J., Kirsche, K., et al. (2018). State-of-the-art technologies for UAV inspections. *IET Radar. Sonar Navig.* 12, 151–164. doi: 10.1049/iet-rsn.2017.0251

Kanaeva, I. A., and Ivanova, J. A. (2021). Road pavement crack detection using deep learning with synthetic data. *IOP Conf. Ser. Mat. Sci. Eng.* 1019, 320–325. doi: 10.1088/1757-899X/1019/1/012036

Khaloo, A., Lattanzi, D., and Cunningham, K., Dell'Andrea, R., and Riley, M. (2018). Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling. *Struct. Infr. Eng.* 14, 124–136. doi: 10.1080/15732479.2017.1330891

Kim, B., Yuvaraj, N., Sri Preethaa, K. R., and Arun Pandian, R. (2021). Surface crack detection using deep learning with shallow CNN architecture for enhanced computation. *Neural. Comput. Appl.* 33, 9289–9305. doi: 10.1007/s00521-021-05690-8

Kong, Y., Yu, Z., Chen, H., Wang, Z., Chen, C., and Guo, B. (2017). "Detecting type and size of road crack with the smartphone" in *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)* (Guangzhou), 572–579.

König, J., Jenkins, M., Mannion, M., Barrie, P., and Morison, G. (2022). What's cracking? A review and analysis of deep learning methods for structural crack segmentation, detection and quantification. *arXiv preprint arXiv:* 2202.03714.

Li, B., Guo, H., Wang, Z., and Li, M. (2022a). Automatic crack classification and segmentation on concrete bridge images using convolutional neural networks and hybrid image processing. *Int. Trans. Infrastruct.* 1, liac016. doi: 10.1093/iti/liac016

Li, G., Ma, B., He, S., Ren, X., and Liu, Q. (2020). Automatic tunnel crack detection based on u-net and a convolutional neural network with alternately updated clique. *Sensors* 20, 717. doi: 10.3390/s20030717

Li, H., Wang, W., Wang, M., Li, L., and Vimlund, V. (2022b). A review of deep learning methods for pixel-level crack detection. *J. Traff. Trans. Eng.* 9, 945–968. doi: 10.1016/j.jtte.2022.11.003

Li, R., Yu, J., Li, F., Yang, R., Wang, Y., Peng, Z., et al. (2023). Automatic bridge crack detection using unmanned aerial vehicle and faster R-CNN. *Constr. Build. Mat.* 362, 129659. doi: 10.1016/j.conbuildmat.2022.129659

Li, W., Huyan, J., Gao, R., Hao, X., Hu, Y., Zhang, Y., et al. (2021). Unsupervised deep learning for road crack classification by fusing convolutional neural network and k_means clustering. *J. Trans. Eng. Pavem.* 147, 04021066. doi: 10.1061/JPEODX.0000322

Li, X., Yang, S., and Liu, H. (2022c). An unsupervised concrete crack detection method based on nnU-Net. *Lectur. Notes Electr. Eng.* 813, 609–623. doi: 10.1007/978-981-16-6963-7_55

Li, Y., Han, Z., Xu, H., Liu, L., Li, X., Zhang, K., et al. (2019). YOLOv3-lite: a lightweight crack detection network for aircraft structure based on depthwise separable convolutions. *Appl. Sci.* 9, 3781. doi: 10.3390/app9183781

Lin, X., Zhang, J., Wu, D., Meng, E., Liu, M., Li, M., et al. (2023). Intelligent identification of pavement cracks based on PSA-Net. *Front. Environ. Sci.* 11, 17. doi: 10.3389/fenvs.2023.1105467

Liu, Y., Yao, J., Lu, X., Xie, R., and Li, L. (2019). DeepCrack: a deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* 338, 139–153. doi: 10.1016/j.neucom.2019.01.036

Lv, X., Duan, F., Jiang, J. J., Fu, X., and Gan, L. (2020). Deep active learning for surface defect detection. *Sensors* 20, 1650. doi: 10.3390/s20061650

Maniat, M., Camp, C. V., and Kashani, A. R. (2021). Deep learning-based visual crack detection using google street view images. *Neural Comput. Appl.* 33, 14565–14582. doi: 10.1007/s00521-021-06098-0

Matarneh, S., Elghaish, F., Al-Ghraibah, A., Abdellatef, E., and Edwards, D. J. (2023). An automatic image processing based on Hough transform algorithm for pavement crack detection and classification. *Smart Sust. Built Environ.* doi: 10.1108/SASBE-01-2023-0004

Mei, Q., and Gui, M. (2020). A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Constr. Build. Mater.* 256, 119397. doi: 10.1016/j.conbuildmat.2020.119397

Mohan, A., and Poobal, S. (2018). Crack detection using image processing: a critical review and analysis. *Alexandria Eng. J.* 57, 787–798. doi: 10.1016/j.aej.2017.01.020

Montero, R., Victores, J. G., Martinez, S., Jardón, A., and Balaguer, C. (2015). Past, present and future of robotic tunnel inspection. *Autom. Constr.* 59, 99–112. doi: 10.1016/j.autcon.2015.02.003

Mubashshira, S., Azam, M. M., and Masudul Ahsan, S. M. (2020). "An unsupervised approach for road surface crack detection," in *2020 IEEE Region 10 Symposium (TENSYMP)*. Piscataway, NJ: IEEE, 1596–1599.

Munawar, H. S., Hammad, A. W., Haddad, A., Soares, C. A. P., and Waller, S. T. (2021a). Image-based crack detection methods: a review. *Infrastructures* 6, 115. doi: 10.3390/infrastructures6080115

Munawar, H. S., Ullah, F., Qayyum, S., and Heravi, A. (2021b). Application of deep learning on uav-based aerial images for flood detection. *Smart Cities* 4, 1220–1242. doi: 10.3390/smartcities4030065

Mundt, M., Majumder, S., Murali, S., Panetsos, P., and Ramesh, V. (2019). "Meta-learning convolutional neural architectures for multi-target concrete defect classification with the concrete defect bridge image dataset," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Piscataway, NJ: IEEE, 11188–11197.

Nguyen, S. D., Tran, T. S., Tran, V. P., Lee, H. J., Piran, M. J., Le, V. P., et al. (2022). Deep learning-based crack detection: a survey. *Int. J. Pavem. Res. Technol.* 16, 943–967. doi: 10.1007/s42947-022-00172-z

Nnolim, U. A. (2020). Automated crack segmentation via saturation channel thresholding, area classification and fusion of modified level set segmentation with Canny edge detection. *Heliyon* 6, e05748. doi: 10.1016/j.heliyon.2020.e05748

Padsumbiya, M., Brahmbhatt, V., and Thakkar, S. P. (2022). Automatic crack detection using convolutional neural network. *J. Soft Comput. Civil Eng.* 6, 1–17. doi: 10.22115/SCCE.2022.325596.1397

Parrany, A. M., and Mirzaei, M. (2022). A new image processing strategy for surface crack identification in building structures under non-uniform illumination. *IET Image Proc.* 16, 407–415. doi: 10.1049/ipr2.12357

Prasanna, P., Dana, K. J., Gucunski, N., Basily, B. B., La, H. M., Lim, R. S., et al. (2014). Automated crack detection on concrete bridges. *IEEE Trans. Autom. Sci. Eng.* 13, 591–599. doi: 10.1109/TASE.2014.2354314

Qu, Z., Cao, C., Liu, L., and Zhou, D. Y. (2021). A deeply supervised convolutional neural network for pavement crack detection with multiscale feature fusion. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 4890–4899. doi: 10.1109/TNNLS.2021.3062070

Qu, Z., Wang, C. Y., Wang, S. Y., and Ju, F. R. (2022). A method of hierarchical feature fusion and connected attention architecture for pavement crack detection. *IEEE Trans. Int. Trans. Syst.* 23, 16038–16047. doi: 10.1109/TITS.2022.3147669

Ren, J., Zhao, G., Ma, Y., Zhao, D., Liu, T., Yan, J., et al. (2022). Automatic pavement crack detection fusing attention mechanism. *Electronics* 11, 3622. doi: 10.3390/electronics11213622

Shi, Y., Cui, L., Qi, Z., Meng, F., and Chen, Z. (2016). Automatic road crack detection using random structured forests. *IEEE Trans. Int. Transp. Syst.* 17, 3434–3445. doi: 10.1109/TITS.2016.2552248

Shim, S., Kim, J., Cho, G.-C., and Lee, S.-W. (2020). Multi-scale and adversarial learning-based semi-supervised semantic segmentation approach for crack detection in concrete structures. *IEEE Access* 8, 170939–170950. doi: 10.1109/ACCESS.2020.3022786

Silva, W. R. L. D., and Lucena, D. S. D. (2018). Concrete cracks detection based on deep learning image classification. *Proc. MDPI* 2, 489. doi: 10.3390/ICEM18-05387

Sony, S., and Laventure, S., and Sadhu, A. (2019). A literature review of next-generation smart sensing technology in structural health monitoring. *Struct. Control Health Monit.* 26, e2321. doi: 10.1002/stc.2321

Sun, M., and Zhao, H., and Li, J. (2022). Road crack detection network under noise based on feature pyramid structure with feature enhancement (road crack detection under noise). *IET Image Proc.* 16, 809–822. doi: 10.1049/ipr2.12388

Sun, Z., Pei, L., Yuan, B., Du, Y., Li, W., Han, Y., et al. (2021). "Pavement crack detection and quantification based on scanning grid and projection method," in *International Conference on Civil Engineering,* Singapore: Springer Singapore, 273–281.

Tabassum, N. (2022). *Image-Based Crack Detection by Extracting Depth of the Crack Using Machine Learning* (Doctoral dissertation). University of South Carolina.

Tang, Y., Qian, Y., and Yang, E. (2022). Weakly supervised convolutional neural network for pavement crack segmentation. *Int. Trans. Infr.* 1, liac013. doi: 10.1093/iti/liac013

Tse, K. W., Pi, R., Sun, Y., Wen, C. Y., and Feng, Y. (2023). A novel real-time autonomous crack inspection system based on unmanned aerial vehicles. *Sensors* 23, 3418. doi: 10.3390/s23073418

Wan, H., Gao, L., Su, M., Sun, Q., and Huang, L. (2021). Attention-based convolutional neural network for pavement crack detection. *Adv. Mat. Sci. Eng.* 2021, 1–13. doi: 10.1155/2021/5520515

Wang, J., He, X., Faming, S., Lu, G., Cong, H., Jiang, Q., et al. (2021). A real-time bridge crack detection method based on an improved inception-resnet-v2 structure. *IEEE Access* 9, 93209–93223. doi: 10.1109/ACCESS.2021.3093210

Wang, L., Zhuang, L., and Zhang, Z. (2019). Automatic detection of rail surface cracks with a superpixel-based data-driven framework. *J. Comput. Civil Eng.* 33, 1–9. doi: 10.1061/(ASCE)CP.1943-5487.0000799

Wu, T., Zhang, H., Liu, J., Chen, Y., Yi, J., Zhang, Z., et al. (2021). "Memory-augment convolutional Autoencoder for unsupervised pavement crack classification," in *2021 China Automation Congress (CAC)*, Piscataway, NJ: IEEE, 2952–2956. doi: 10.1109/CAC53003.2021.9727812

Xiang, X., Zhang, Y., and El Saddik, A. (2020). Pavement crack detection network based on pyramid structure and attention mechanism. *IET Image Proc.* 14, 1580–1586. doi: 10.1049/iet-ipr.2019.0973

Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., Ling, H., et al. (2019). Feature pyramid and hierarchical boosting network for pavement crack detection. *IEEE Trans. Int. Transp. Syst.* 21, 1525–1535. doi: 10.1109/TITS.2019.2910595

Yang, N., Li, Y., and Ma, R. (2022). An efficient method for detecting asphalt pavement cracks and sealed cracks based on a deep data-driven model. *Appl. Sci.* 12, 10089. doi: 10.3390/app121910089

Yang, Z., Ni, C., Li, L., Luo, W., and Qin, Y. (2022). Three-stage pavement crack localization and segmentation algorithm based on digital image processing and deep learning techniques. *Sensors* 22, 8459. doi: 10.3390/s22218459

Yu, G., Dong, J., Wang, Y., and Zhou, X. (2022). RUC-Net: a residual-Unet-based convolutional neural network for pixel-level pavement crack segmentation. *Sensors* 23, 53. doi: 10.3390/s23010053

Yuan, G., Li, J., Meng, X., and Li, Y. (2022). CurSeg: a pavement crack detector based on a deep hierarchical feature learning segmentation framework. *IET Intellig.Transp. Syst.* 16, 782–799. doi: 10.1049/itr2.12173

Zhang, L., Yang, F., Zhang, Y. D., and Zhu, Y. J. (2016). "Road crack detection using deep convolutional neural network," in *2016 IEEE international conference on image processing (ICIP),* Piscataway, NJ: IEEE, 3708–3712.

Zhao, M., Shi, P., Xu, X., Xu, X., Liu, W., Yang, H., et al. (2022). Improving the accuracy of an R-CNN-based crack identification system using different preprocessing algorithms. *Sensors* 22, 7089. doi: 10.3390/s22187089

Zhong, K., Teng, S., Liu, G., Chen, G., and Cui, F. (2020). Structural damage features extracted by convolutional neural networks from mode shapes. *Appl. Sci.* 10, 4247. doi: 10.3390/app10124247

Zhong, M., Sui, L., Wang, Z., and Hu, D. (2020). Pavement crack detection from mobile laser scanning point clouds using a time grid. *Sensors* 20, 4198. doi: 10.3390/s20154198

Zhu, J., and Song, J. (2020). Weakly supervised network based intelligent identification of cracks in asphalt concrete bridge deck. *Alexandria Eng. J.* 59, 1307–1317. doi: 10.1016/j.aej.2020.02.027

Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., Wang, S., et al. (2018). Deepcrack: learning hierarchical convolutional features for crack detection. *IEEE Trans. Image Proc.* 28, 1498–1512. doi: 10.1109/TIP.2018.2878966

frontiers | Frontiers in Environmental Science

TYPE Review
PUBLISHED 20 February 2024
DOI 10.3389/fenvs.2024.1336088

OPEN ACCESS

EDITED BY
Chengye Zhang,
China University of Mining and Technology, China

REVIEWED BY
Shashi Kant Gupta,
Eudoxia Research University, United States
Ioannis G. Tsoulos,
University of Ioannina, Greece
Wu Ling,
China University of Geosciences, China
Jun Yue,
Central South University, China

*CORRESPONDENCE
Yong Suk Chung,
yschung@jejunu.ac.kr
Sheikh Mansoor,
mansoorshafi21@gmail.com

RECEIVED 10 November 2023
ACCEPTED 06 February 2024
PUBLISHED 20 February 2024

CITATION
Popescu SM, Mansoor S, Wani OA, Kumar SS, Sharma V, Sharma A, Arya VM, Kirkham MB, Hou D, Bolan N and Chung YS (2024), Artificial intelligence and IoT driven technologies for environmental pollution monitoring and management.
Front. Environ. Sci. 12:1336088.
doi: 10.3389/fenvs.2024.1336088

Artificial intelligence and IoT driven technologies for environmental pollution monitoring and management

Simona Mariana Popescu[1], Sheikh Mansoor[2]*, Owais Ali Wani[3], Shamal Shasang Kumar[3], Vikas Sharma[4], Arpita Sharma[5], Vivak M. Arya[4], M. B. Kirkham[6], Deyi Hou[7], Nanthi Bolan[8] and Yong Suk Chung[2]*

[1]Department of Biology and Environmental Engineering, University of Craiova, Craiova, Romania, [2]Department of Plant Resources and Environment, Jeju National University, Jeju-si, Republic of Korea, [3]Division of Soil Science and Agricultural Chemistry, Sher-e-Kashmir University of Agricultural Sciences and Technology, Sopore, Jammu and Kashmir, India, [4]Division of Soil Science and Agricultural Chemistry, Faculty of Agriculture, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Chatha, Jammu and Kashmir, India, [5]Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL, United States, [6]Department of Agronomy, Kansas State University, Manhattan, KS, United States, [7]Soil and Groundwater Division, School of Environment, Tsinghua University Beijing, China, [8]UWA School of Agriculture and Environment, Institute of Agriculture, Perth, WA, Australia

Detecting hazardous substances in the environment is crucial for protecting human wellbeing and ecosystems. As technology continues to advance, artificial intelligence (AI) has emerged as a promising tool for creating sensors that can effectively detect and analyze these hazardous substances. The increasing advancements in information technology have led to a growing interest in utilizing this technology for environmental pollution detection. AI-driven sensor systems, AI and Internet of Things (IoT) can be efficiently used for environmental monitoring, such as those for detecting air pollutants, water contaminants, and soil toxins. With the increasing concerns about the detrimental impact of legacy and emerging hazardous substances on ecosystems and human health, it is necessary to develop advanced monitoring systems that can efficiently detect, analyze, and respond to potential risks. Therefore, this review aims to explore recent advancements in using AI, sensors and IOTs for environmental pollution monitoring, taking into account the complexities of predicting and tracking pollution changes due to the dynamic nature of the environment. Integrating machine learning (ML) methods has the potential to revolutionize environmental science, but it also poses challenges. Important considerations include balancing model performance and interpretability, understanding ML model requirements, selecting appropriate models, and addressing concerns related to data sharing. Through examining these issues, this study seeks to highlight the latest trends in leveraging AI and IOT for environmental pollution monitoring.

KEYWORDS

environmental pollution, monitoring, safety, advance technologies, machine learning, IoT

1 Introduction

Hazardous substances in the environment are those that pose a threat to human health, plant, and animal life, or the environment. These substances include heavy metals, pesticides, herbicides, and persistent organic pollutants (POPs) that have been introduced into the environment through various means (Young et al., 2004). Sources of hazardous substances in soil include industrial activities, improper disposal of hazardous waste, agricultural practices, and natural processes such as erosion and weathering. These substances can persist in the environment for long periods and can have a negative impact on soil quality, plant growth, and human health (Bachmann, 2006; Baran et al., 2011; Bolan et al., 2021; Rani et al., 2021). The effects of hazardous substances in the environment can vary depending on the type and concentration of the substance, as well as the duration of exposure. Some hazardous substances can cause acute health effects, such as respiratory problems, skin irritation, poisoning, nausea, and vomiting, while others can lead to chronic health problems, including cancer, reproductive disorders, and developmental abnormalities (Baran et al., 2011; Li et al., 2022; Yang et al., 2022).

Effective management of hazardous substances in the environment requires monitoring, remediation, and prevention strategies. Monitoring involves regular testing of soil for the presence of hazardous substances, which allows for early detection and appropriate management strategies to be implemented. Remediation involves the removal or treatment of contaminated sites. Prevention strategies include reducing the use of hazardous materials (i.e., source control) and implementing best practices for waste disposal and land use. Effective management is essential for ensuring the long-term health and sustainability of soil and its ecosystems (Mansoor et al., 2022; Sharma et al., 2022; Sonne et al., 2023).

Real-time monitoring of hazardous materials in soil and plants is an important task that can help to ensure the safety of food crops, protect the environment, and prevent human exposure to harmful substances. The use of artificial intelligence (AI) powered sensors and devices can greatly enhance the accuracy and efficiency of this monitoring process. AI-powered sensors and devices can be used to detect and quantify the presence of various hazardous materials in soil and plants (Wilson, 2012; Yang et al., 2021). These sensors and devices can be designed to measure parameters such as pH, temperature, moisture, conductivity, and various chemical properties of the soil and plant tissue. Machine learning algorithms can be used to analyze the data collected by these sensors and devices, enabling the identification of specific hazardous materials in real-time (Wilson, 2012; Yang et al., 2021). These algorithms can also be used to predict the potential impact of these materials on human health and the environment. For example, E-nose (olfactory) algorithms are used to analyze data generated by sensors and identify the presence of hazardous chemicals based on their unique chemical signature (Jeong and Choi, 2022). These algorithms can use a variety of techniques, such as pattern recognition, artificial neural networks, and fuzzy logic. One of the key advantages of E-nose technologies is their ability to detect hazardous chemicals in real-time, allowing for immediate response to potential threats. E-nose technologies can be used for a variety of applications, such as monitoring air quality in urban areas, detecting leaks from industrial processes, and detecting explosives and other hazardous materials (Jeong and Choi, 2022).

The use of AI-powered sensors and devices for real-time monitoring of hazardous materials in soil and plants has several benefits (Singh and Kaur, 2022). Firstly, it allows for more accurate and reliable detection of these materials compared to traditional laboratory-based methods. Secondly, it provides real-time data, allowing for quick responses to any potential contamination events. Finally, it reduces the need for manual data collection and analysis, reducing the workload and increasing the efficiency of the monitoring process (Jeong and Choi, 2022). The use of AI-powered sensors and devices for real-time monitoring of hazardous materials in soil and plants is a promising approach that can help to ensure the safety of food crops, protect the environment, and prevent human exposure to harmful substances. Various approaches can be used for AI-based toxicity prediction, including machine learning methods, deep learning methods, and hybrid approaches that combine both methods. Integrating various sources of data, such as chemical structures, toxicological and physiological data, and environmental factors, to improve the accuracy and reliability of toxicity predictions, is important (Jeong and Choi, 2022; Chen et al., 2023).

One of the recent advances is the combination of AI and Internet of Things (IoT) technologies, for particulate matter (PM) monitoring, which uses low-cost sensors that can be easily deployed in various environments (Bhagat et al., 2020). These sensors can collect data on PM levels and send it to a centralized platform for analysis. AI algorithms can then process these data to provide real-time information on PM levels and predict future trends. Heavy metals can also be monitored using AI. Numerous studies have been conducted over the last 10 years to forecast the effectiveness of heavy metal removal from soil using machine learning (Zafar et al., 2017; Zhu et al., 2019). AI models for the optimization and prediction of heavy metal removal include black box, fuzzy logic, kernel, evolutionary, and hybrid models.

With the rapid changes that the environment is experiencing, data sharing and reuse with the help on AI algorithms and instruments (Shen, 2018), plays an important role in supporting researchers to safeguard the continuous threatened environment and ensure the implementation of sustainable environmental management practices (Aggestam and Mangalagiu, 2020). Scientists can make use of online data sharing tools and platforms that comprise vast and intricate Earth and environmental science data like climatic and atmospheric data, pedology, hidrology, ecology, and biodiversity data (Crystal-Ornelas et al., 2022; Basel et al., 2023) for testing, analyzing, interpretation of theories, prediction models and experimental data (Kostal et al., 2022) that lead to better undestanding environmental issues.

In this comprehensive review, we will examine the application of artificial intelligence (AI) in monitoring hazardous materials across different environments, namely, soil, air, and water. We will explore the latest breakthroughs and progress made in this field, which integrates machine learning algorithms with sensor technologies. We will also consider the benefits and drawbacks associated with AI-powered monitoring systems. This new field has revolutionized soil, air, and water monitoring, enhancing accuracy, efficiency, and timeliness in detecting and analyzing hazardous substances.

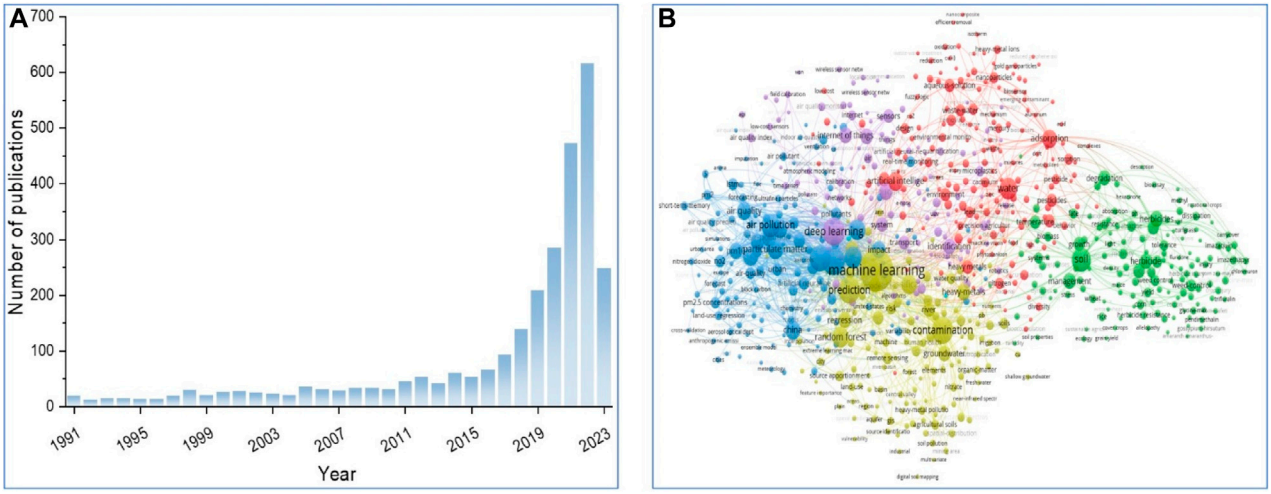


FIGURE 1
Systematic literature search results on AI-Driven Technologies for Hazardous Substance Monitoring topic. **(A)** The number of publications related to AI-Driven Technologies for Hazardous Substance Monitoring topic in each year following literature screening. **(B)** A network visualization map delineating keywords within the associated literature is depicted, revealing clusters indicative of distinct research themes. Nodes sharing analogous colors denote cohesive clusters comprising interrelated terms. This map was generated through the utilization of VOSviewer.

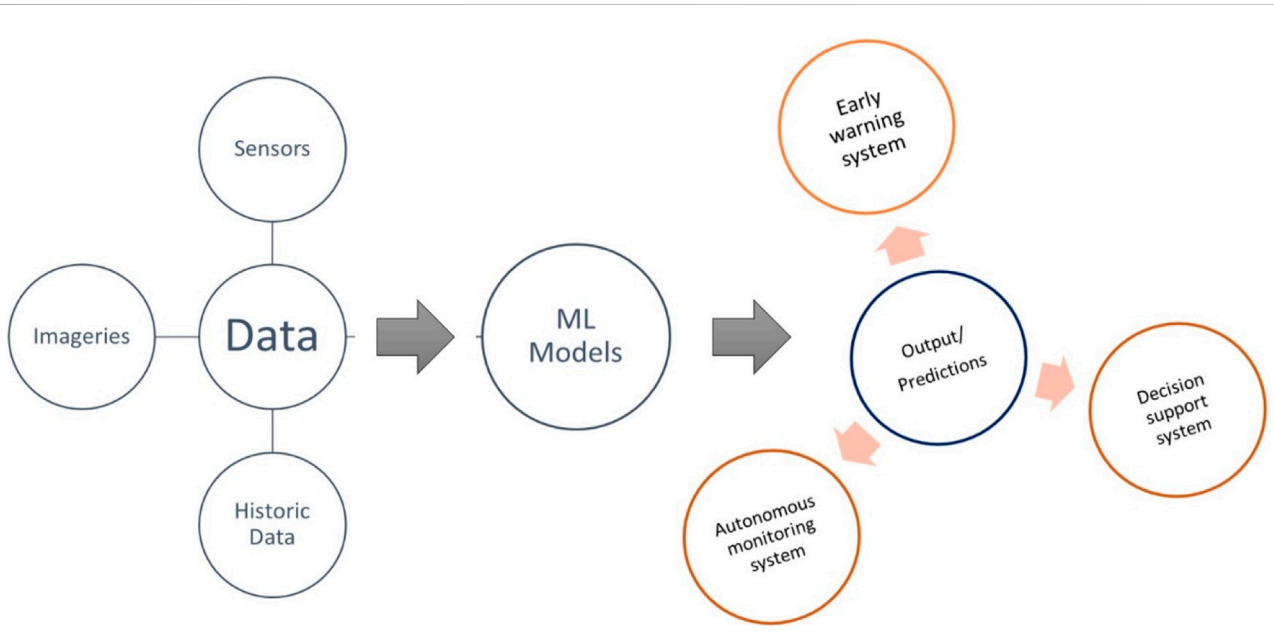


FIGURE 2
AI solutions: making use of Data and AI algorithms for hazardous substance monitoring and containment.

A literature search was conducted in Web of Science Core Collections with the following search terms (TS stands for Topic Searches): TS=("artificial intelligence" or "AI" or "machine learning" or "deep learning" or "Internet of Things" or "IoT" or "computer vision" or "robotics" or "natural language process" or "real-time monitoring" or "e-nose") AND TS=("hazardous substance" or "hazardous chemical" or "hazardous material" or "pollutant" or "contamination" or "toxin" or "heavy metal" or "pesticide" or "herbicide" or "persistent organic pollutants" or "POPs" or "microplastic") AND TS=("environment" or "soil" or "terrestrial" or "aquatic" or "aqueous" or "freshwater" or "lake" or "river" or "sediment" or "marine" or "ocean" or "air" or "atmosphere"). A total of 2828 results were retrieved. The results were visualized using the VOS (visualization of similarities) viewer software (version 1.6.19). Figure 1 presents the systematic literature search results covering the number of publications and the keyword co-occurrence map on the topic of the review (AI-Driven Technologies for Hazardous Substance Monitoring).

2 AI solutions for hazardous substance monitoring in different environments

In recent years, there has been an increase in interest for using AI to anticipate and predict environmental pollution. We can split AI solutions into three steps or phases that include inputs (data), models (AI algorithms), and outputs (monitoring or decision support) (Figure 2). Data forms the basis of any AI solution. These models work best when they have a high number of data points, especially the ones that are coming from environments that are continually being monitored and expending solutions for probable actions. Data sources today can be a variety of sensors, ranging from imaging to non-imaging types or remote to in-contact sensors that provide large volumes of data. Then there are the historic or legacy data. AI algorithms can analyse massive amounts of sensor readings, historical data, and other important information from monitoring systems. AI can detect hazardous material levels and contamination events by identifying patterns, trends, and anomalies in data using machine learning and data mining. Image analysis can detect hazardous material spills via satellite photos or drone images. Computer vision algorithms can recognise chemicals, vegetation changes, and pollution sources.

AI algorithms can be trained on past data to predict hazardous material releases and environmental pollution. AI solutions for environmental monitoring, thus, would include early warning systems for hazardous material release, autonomous pollution monitoring systems as well as decision support systems. These models can help authorities and organisations prepare for and respond to emergencies. These models have proved useful for multiple environmental conditions, be it soil, air, or water. AI modellers, however, should offer sufficient details to explain and support the selection of model parameters, as well as their creation and assessment.

Many short- and long-term forecasting applications use ANNs (artificial neural networks). One has created an IoT-enabled environmental toxicology model to detect air pollution (Asha et al., 2022). The model uses artificial intelligence to report the status of the quality of air in real-time utilising a cloud server and broadcasts alarms when hazardous pollutants are present. The AAA (artificial algae algorithm)-based ENN (Elman neural network) model classifies and predicts air quality in future timestamps (a timestamp is the current time of an event that a computer records). WiFi gateways send the data collected by sensors to a cloud server. AAA optimises the ENN model parameters during data processing. For monitoring and enforcement, mobile electronic-nose (E-nose) devices and algorithms have been developed to quickly detect pollution from point sources. Due to their sensitivity, E-nose devices can detect carbon, GHG emissions, and particle pollutants emitted into air or water (Wilson, 2012).

Traditional soil mapping entails physically collecting soil samples and transferring them to a laboratory for additional analysis (Signes-Pastor et al., 2016; Sharma et al., 2017). The advent of technologies permitting high-resolution, quick, and inexpensive mapping of soil pollutants has favourable aspects over traditional techniques. Jia et al. (2021) created an unique modelling method that forecasts soil arsenic levels using high resolution aerial imagery (HRAI) photos. The method makes use of cameras that are installed on aircraft to take high-resolution (0.1–0.5 m) pictures of broad areas. The first layer of a model for displaying soil arsenic levels shows a thorough report on soil contamination and HRAI. The image is broken down into pixels in the second layer, and sample points are represented by pixel features. In order to forecast the risk levels of arsenic, four alternative machine learning algorithms were constructed. The Extreme Random Forest (ERF) algorithm had the best prediction and accuracy (Jia et al., 2021). Remote sensing and aerial imageries provide continuous spatial data which, coupled with machine learning models, are providing highly accurate maps of hazardous substances in the environment that were not possible with standard geostatistical techniques.

Microplastics in the environment have become a cause for concern, but their evaluation in soils is a laborious process. Hyperspectral imaging (HSI) and convolutional neural network (CNN) technology-based methods have been developed for identifying microplastic polymers (Ai et al., 2023). The technique developed by Ai et al. (2023) provides a non-destructive, rapid approach to detection of microplastics in soils. Microfluidic devices using machine learning and AI promise to be next-generation monitoring systems (Pouyanfar et al., 2022). Sensitive microfluidic devices produce high-quality pollutant data and reveal important environmental information. Artificial intelligence can categorize, characterize, and predict data from microfluidic systems. The two can be easily set up. Ahmadi et al. (2021) used ANN to predict the concentrations of organophosphorus pesticides in water. Their model worked best for prediction of Malathion ($R^2 = 0.887$), followed by parathion ($R^2 = 0.711$), and Diazinon ($R^2 = 0.714$). However, the R-squared value was acceptable for all pesticide types (Sarkar and Pandey, 2015). Used flow discharge data from the Yamuna River in northern India, as well as biochemical oxygen demand, temperature, pH, and dissolved oxygen collected on a monthly basis in the vicinity of Mathura, India, and employed a feed forward, error back propagation algorithm to develop ANN models. The predictions from the ANN models were accurate, with correlation values being as high as 0.9, indicating that ANNs can be efficiently utilized to predict water quality.

3 AI-powered technologies for pollution monitoring

3.1 Spectroscopy

Spectroscopy, which involves measuring the functions of energy with matter, is of preeminent importance in remote sensing. It has been widely utilized in field of chemistry and astronomy to identify materials, and advancements in instrumentation have led to its increasing use in remote sensing studies (Slonecker et al., 2010). Visible and near-infrared reflectance spectroscopy is an environmentally friendly and cost-efficient technique that shows promise for estimating concentrations of various heavy metals in soil. Additionally, it offers a viable alternative for assessing heavy metal levels across large areas and for an extended period (Shi et al., 2022).

For example, Zhao et al. (2022) used visible and near-infrared spectroscopy with extreme gradient boosting (XGBoost) - as an

effective machine learning technique to create an estimation model for heavy metal pollution in mangrove sediment sites. Vis-NIR spectroscopy coupled with partial least squares (PLS) and radial basis function neural network (RBFNN) prediction models, were engaged by Sanaeifar et al. (2022) in their study to investigate the effects of airborne Pb on tea plants. Elevated concentrations of Pb had deleterious effects on the plant and the results revealed PLS–RBFNN models to be more accurate and superior to conventional methods in terms of prediction, giving Vis-NIR spectral data traits like high speed and simplicity in monitoring heavy metal pollution.

Lately, there has been significant interest in the development and application of flexible, surface-enhanced Raman scattering (SERS) substrates for the detection of hazardous substances. Bharati and Soma (2021) conducted a comprehensive 4-year investigation into various flexible SERS substrates, such as paper or cellulose, polymer nanofibers, 3D sponges, and fabrics. They explored the potential of these substrates for on-site detection of explosives, pesticides, chemical warfare agents, and drugs, and the research was related to fields such as homeland security, food safety, and medicine. The results of their study highlight the considerable opportunities for SERS substrates in the detection and monitoring of hazardous waste.

3.2 Ground-based monitoring sensors

As reported by Pant et al. (Pant et al., 2019), air quality monitoring mainly employs two types of sensor networks: manual and automatic monitoring sensors (like wireless or community sensor networks). According to Kim et al. (Kim et al., 2008; Kumar et al., 2022), after combining air quality monitoring sensors with geo-sensor network technologies, the collection and interpretation of geospatial and pollution data were transformed. Levy et al. (2010) observed that the monitoring of pollutants, like particulate matter, nitrite, and sulphur dioxide, can be achieved using various sensor networks, including instruments like Met One Instrument BAM-1020 Beta Attenuation Monitor (Met one instruments ltd. 1,600 Washington Blvd. Grants Pass, Oregon), Alphasense OPC-N2 Particle Monitor (Halo labs 1828 Burlingame, CA, 94,010 USA) and Aeroqual Series 500 with NO_2 Sensor Head (Aeroqual Limited, London). Ground-based aerosol optical measurements not only help characterize ambient aerosols, but also assist in validating satellite retrievals and numerical modeling algorithms (Levy et al., 2010; Mansoor et al., 2021). Amado and Cruz (Amado et al., 2018; Naz et al., 2023) used machine learning to calibrate a predictive model for monitoring and characterizing air quality. Their methodology involved creating a prototype with integrated sensors, including DHT 11 temperature and relative humidity sensors, as well as MQ2, MQ5, and MQ135 gas sensors (Zhengzhou Winsen Electronics Technology Co., Ltd.). The study developed five predictive models: k-nearest neighbors (KNN), support vector machine (SVM), naïve-Bayesian classifier, random forest, and neural network. The models demonstrated 99% accuracy.

Artificial intelligence has found its way into agricultural systems, incorporating various sensors for improved performance. Soil moisture sensors are used to ensure adequate irrigation, while temperature sensors monitor the ambient conditions of grain storage. Manickavasagan et al. (2006) utilized temperature and humidity monitoring technology to enhance quality control of stored products by indirectly measuring key parameters. This approach also enabled the prediction of quantitative and qualitative losses in stored grains, providing valuable insights for decision-making in agricultural businesses.

Under field conditions, Liu et al. (2019) explored the application of acoustic sensors, specifically piezoelectric sensors, for detecting insect infestation. Piezoelectric crystals, which transform mechanical stress into electrical charge, are the fundamental component of an acoustic sensor. By measuring sound waves or vibrations and transforming them into electronic signals, piezoelectric sensors can effectively indicate the presence of insects. To aid decision-making and improve management practices, Cruz et al. (2018) conducted a study on predicting flood levels in advance. They put in place a real-time monitoring system with numerous sensors that can gauge variables including rainfall volume and intensity, soil moisture, water level, and rate of water level rise. Using a multi-layered artificial neural network developed with MATLAB, they created a prediction model based on the gathered sensor data. The results showed a high level of accuracy, with goodness-of-fit values of 0.99889 for the training dataset, 0.99362 for the test dataset, and 0.99764 for the validation dataset.

3.3 Aerial imaging and unmanned aerial vehicles (UAVs)

Since the early 1990s, aerial imagery has been employed to monitor hazardous waste. Aerial photographs have proven useful in various applications for detecting and analyzing the presence of hazardous waste, waste-disposal sites, and landfills (Pope et al., 1996; Wani et al., 2023). Historical aerial photographs provide valuable documentation for compiling a record of changes occurring at hazardous sites and are considered reliable for monitoring changes over time. The U.S. Environmental Protection Agency (USEPA) has made use of an aerial imagery library that dates back to the 1930s to retrace the history of waste management and disposal at hazardous waste sites. Environmental clean-up programmes have generated more than 4,000 historical aerial photographic reports on hazardous waste activity and employed them in clean-up (Benger et al., 2004).

Aerial photography allows for the interpretation of various objects at hazardous waste sites, including evidence of discarded materials, barrels and drums, open dumps, spills, and disturbance. These features enable monitoring and analysis of potential impacts related to hazardous waste. Aerial imagery facilitates the identification of vegetation patterns, investigation of local groundwater movement to assess potential pollutant migration, determination of drainage routes, examination of hydrological conditions, and evaluation of subsequent land use on closed landfills (Slonecker et al., 2002; Garofalo, 2003).

Hyperspectral imagery-based models have proven effective in predicting heavy-metal distribution in soils. In a study conducted by Tan et al. (Tan et al., 2021), a competitive adaptive reweighted sampling (CARS) method was proposed for this purpose. The researchers compared the accuracy of various models and

discovered that CARS in combination with a stacking method exhibited the highest accuracy and stability. This method utilized the spectrum in the range of 2–2.3 μm, which is a common characteristic band for heavy metals. Additionally, it effectively addressed challenges, such as overfitting due to imbalanced data and limited training sample sets. Importantly, even in the presence of spatial heterogeneity, the distribution of heavy-metal concentrations derived from the CARS-stacking method showed consistency in the verification analysis.

Increased population and industrialization have led to a rise in hazardous waste spillage or leakage incidents. Detecting and monitoring toxic, flammable, and inert gas leaks are a global priority and have been the subject of extensive research. The emergence of Unmanned Aerial Vehicles (UAVs) or drones has significantly enhanced the collection of aerial imagery. This technological advancement has provided a boost in obtaining high-resolution and up-to-date aerial data for various applications, including the monitoring and analysis of heavy-metal distribution in soils. Forward reconnaissance drones, equipped with sensors, are valuable tools for enhancing situational awareness in environments that are challenging for humans. Remote gas detection systems, utilizing mobile robotic platforms such as drones, have emerged as a promising approach in this field, especially in environments too hazardous for human exploration (Gerhardt et al., 2014).

By augmenting drones with sensors, hazardous materials or spills can be identified, and waste can be monitored to enable effective planning and management, thereby minimizing the exposure of first-response teams (Seiber et al., 2018). Also, integrating sensor-based particulate detection with autonomous drone flight control enables dynamic identification and real-time tracking of aerial plume boundaries. The findings of Seiber et al. (2018) demonstrate that UAVs can precisely recognize and track contaminant plumes over time, thus providing visual indicators and aid data collection that can be used to validate and advance the plume movement models. In synopsis, the utilization of sensor-equipped drones for forward reconnaissance purposes can significantly contribute to identifying and monitoring hazardous materials, spills, and gas leaks in environments where human access is unsafe. The integration of sensors and autonomous flight control enhances the capabilities of drones in real-time plume tracking and provides valuable data for improved modelling and management of hazardous incidents.

In incidents that involve hazardous materials, drone flights play a crucial role in promptly and accurately identifying the direction of spillage or gaseous material dispersion (Restas, 2015). Numerous studies have highlighted the ability of UAVs to be rapidly deployed to disaster sites and gather essential data about the extent and impact of spills. This information is vital for effective planning, management, and response efforts. Integrating UAVs into the toolkits of industry and government spill response teams can greatly enhance response capabilities, mitigation strategies, and overall accountability (Messinger and Silman, 2016). In summary, UAVs are indispensable in assessing and managing incidents involving hazardous materials. Their ability to be quickly deployed, gather data, and monitor critical areas provides valuable insights for response teams, enabling them to make informed decisions, mitigate risks, and safeguard both human lives and the environment.

3.4 Ground robotics

In recent years, ground robotics have emerged as a technology with gas-sensitive sensors that can be employed to assess situations in enclosed and unventilated spaces by detecting gas distributions, essentially acting as a "sense of smell." A study conducted by Wandel et al. (Wandel et al., 2003) aimed to develop a unified algorithm for localizing the source of odors using ground robotics. They conducted experiments with ethanol as a test substance and successfully determined its concentration. However, a few irregularities occurred due to external factors that influenced the data. For instance, summertime natural convection caused heat transfer from the glass window in the enclosed space, increasing the concentration of ethanol. The researchers demonstrated the advantage of this tool in its ability to monitor suspected areas and detect the presence of hazardous substances.

Vincent et al. (2019) conducted an experiment in a laboratory wind tunnel and a real-world environment to study the formation of gas plumes. Typically, sensors used for gas detection are characterized using data from controlled gas rigs, where precise step changes in gas mixtures are produced. However, in real-world scenarios, such as mobile robot exploration, gas sources do not emit steady concentrations. Instead, they generate plumes, resulting in areas of high and low gas concentrations in the surrounding environment.

An algorithm method to control mobile robots - iRobot Create was designed by Ahmed et al. (2016). The robotized mobile nodes are integrated in a WSN structure, have built-in sensors, equipped with a microprocessor Gumstix Verdex Pro™ XL6P. The real-time experiments demonstrated the suitability of this wireless sensor network robotized nodes that enables two robots to receive information from each other for detecting a pollutea area, based on pollution searching, reorientation and surveillance parameters.

Haldorai et al. (2024) developed a robot for polluted water surface cleaning. The autonomus robot is equiped with two Arduino microcontrollers, powered by a 6 V lead-acid battery, implemented with SSD real-time object detection deep learning model for the precise detection of wastes from water surface. The robot reaches maximum potential, can avoid obstacles and colect by itself the wastes from the water surface when it gets to 45 and 75 cm from the floating debris making it an effective tool for waste removal from water bodies.

As highlighted by Tsitsimpelis et al. (2019), robotic systems offer an ideal solution for identifying and monitoring extreme radiation exposure levels, as well as toxic and combustible atmospheres. These systems address several challenges by eliminating the need for humans to physically access hazardous locations. Furthermore, they can provide valuable data on the conditions of these places that would otherwise remain inaccessible. Ground robotics equipped with gas-sensitive sensors offer a valuable means of monitoring and detecting hazardous materials. Despite certain challenges and external influences on data accuracy, this technology holds promise in identifying and localizing odour

sources in enclosed spaces, contributing to the overall effort of managing and minimizing the presence of dangerous substances.

3.5 Satellite remote sensing

Remote sensing is a scientific and technological approach that enables the identification and assessment of various characteristics and properties of Earth's targets from a distance. On both a global and local level, it has offered systematic and repeated observations of numerous features of the Earth's surface, including the atmosphere, water, land, living things, vegetation, pollution, and climate. The utilization of remote sensing has played a crucial role in detecting and quantifying pollution rates, mapping, monitoring, and mitigating pollution (Mertikas et al., 2021).

One of many applications of remote sensing is monitoring and managing hazardous waste and sites. Multispectral sensors (MSS) mounted on remote sensing platforms are capable of digitally collecting energy levels of reflectance in specific bands of the electromagnetic spectrum. These systems offer advantages such as statistical analysis of data and the ability to extend observations beyond the capabilities of aerial photography. Land use, regional risk assessment, and hazardous waste site spectral characteristics and pollution profiles have been monitored using multispectral imaging systems mounted on satellites and various aircraft-based systems. For instance, Bølviken et al. (Bølviken et al., 1977) demonstrated that MSS data could be employed to identify heavy metal contamination based on fundamental spectral characteristics.

Furthermore, multispectral imagery has been used by several researchers to identify and locate previously unknown as well as illegal hazardous waste sites. Errico et al. (Errico et al., 2014) proposed a methodology that combines synthetic aperture radar (SAR), multispectral data, and GIS-based processing for detecting environmental hazards. This system yielded satisfactory results and contributed to countering and controlling environmental crimes. Additionally, hyperspectral imagery can be employed to identify and map the spatial distribution of various heavy metals. For instance, Kemper and Sommer (Kemper et al., 2004) used a HyMap sensor and utilized airborne hyperspectral images to map lead and arsenic contamination in the Guadiamar flood plain, Andalusia. Similarly, Wu et al. (2011) reported satisfactory outcomes in heavy metal mapping in Nanjing City of China, using simulated HyMap data. They reported that direct predictions based on hyperspectral images often require signals from bare soils, which can be achieved during winter or early spring or when there is low vegetation coverage due to crop rotation in agricultural areas. Therefore, remote sensing has been instrumental in monitoring and managing hazardous waste and sites. Multispectral sensors and airborne hyperspectral images have provided valuable data for assessing pollution, identifying contamination, and mapping the distribution of heavy metals. These technologies offer a non-intrusive and comprehensive approach to understanding and addressing environmental concerns related to hazardous waste.

The incorporation of diverse advanced technologies has significantly improved the field of environmental monitoring, especially in the identification and control of hazardous waste. Spectroscopy, rooted in chemistry and astronomy, has been widely applied in remote sensing. Visible and near-infrared reflectance spectroscopy is presented as an environmentally friendly and cost-effective technique for approximating concentrations of heavy metals in soil. Machine learning methods, such as extreme gradient boosting and neural networks, enhance the accuracy of predictions regarding heavy metal pollution when combined with spectroscopic data. Ground-based monitoring sensors, encompassing both manual and automatic networks and utilizing machine learning models, play a role in robust air quality monitoring. Aerial imaging, covering historical images and hyperspectral imagery, provides a comprehensive perspective on hazardous waste sites, aiding in the evaluation of pollution. Unmanned Aerial Vehicles (UAVs) equipped with sensors are essential for real-time identification and tracking of contaminant plumes, thereby improving response capabilities. Ground robotics, equipped with gas-sensitive sensors, are valuable in enclosed spaces for the detection of hazardous substances. Ultimately, remote sensing, facilitated by multispectral and hyperspectral sensors, serves as a potent tool for systematic Earth observation, particularly in the identification of heavy metal contamination and mapping the distribution of pollutants. This holistic approach underscores the varied strategies employed to tackle environmental issues and emphasizes the pivotal role of advanced technologies in preserving ecosystems.

4 Maximizing safety in hazardous environments with AI-Driven monitoring

Nowadays awareness of the safety of employees working in the production and manufacturing operations in different industrial environments that involve hazardous substances and materials has become of utmost interest (Farrokhi-Asl et al., 2020). Thus, monitoring the concentration and leakage of pollutants (Fung et al., 2019) and development of safe and efficient methods (Fung et al., 2019) to reduce the exposure of humans and the detrimental risks for the environment, with fewer accidents provoked by hazardous substances (Wong et al., 2018; Wang et al., 2020), are becoming a priority in the implementation of safety plan management in many sectors of the economy (Binajjaj et al., 2018). In pollutant discharge events, immediate, precise, and intelligent intervention is needed to alarm, prevent, and control hazardous leakage (Mendil et al., 2022; Wang B. et al., 2023). Due to the fact that humans cannot identify in time possible threats, because the majority of gases are odorless, colorless, and tasteless (Visvanathan et al., 2018), implementation in environmental monitoring and detection of artificial intelligence based devices is getting more popular (Daam et al., 2019; Emaminejad and Akhavian, 2022). The popularity of AI and unmanned machines arises from the effective, real-time, automated solutions that these technologies provide when placed in hazardous environments. They can detect and isolate possible threats before they cause harm, with less or no human involvement (Palacín et al., 2019; Das et al., 2020; Jiang et al., 2022). Recent advances have been developed for the combination of autonomous robots with sensors that are flexible and easily and remotely deployed in unsafe and toxic environments (Ristic et al., 2017). They can improve the safety of workers and

neighborhood areas without the risk of involving human life (Fan et al., 2019). These devices are embedded with systems capable of exploring, monitoring, detecting, and alarming, in case hazardous events occur (Maedche et al., 2019). For example, Joshna et al. (Joshna et al., 2019) developed an independent robot using Arduino UNO (R3), equipped with sensors that can detect and identify toxic gases. It can also perform degasification making use of oxidizing agents that are sprayed on the gases, reducing their deleterious effects. Fan et al. (Fan et al., 2019) developed a mobile robotic system equipped with an E-nose that is highly accurate in detecting and identifying different toxic gases like CO and NO_2 present in the environment. It can be used in fire and emergency response departments. Gallego et al. (2015) developed an unmanned aerial vehicle (UAV) equipped with metal oxide semiconductor (MOX) gas sensors, capable of detection of toxic gas leaks, to monitor gas pipes in outdoor areas. Novelty traits of the system include improved velocity and reduced energy consumption and investment costs.

A wireless microcontroller (MCU) offers acquisitions of real-time data, and a GPS/GSM (Global Positioning System/Global System for Mobile Communication) modem offers accurate registration and communication of the location and route in addition to measuring the gas concentration. With the intention of maximizing safety for staff workers in factories, Shi et al. (2016) Burgués et al. (2019) proposed a Crazyflie 2.0 nano-drone assembled with a MOX gas sensor. This nano-air vehicle (NAV) can accurately localize the gas source with instantaneous gas distribution mapping, which requires less time for concentration measurements than other previous devices. Similarly, Das et al. (Das et al., 2020) proposed a robot that can be deployed in an unknown and uneven environment that can recognize hazardous gases (carbon dioxide, vaporized alcohol, and liquefied petroleum gas) with an average accuracy of 98%. The robot can be operated to avoid collision obstacles and detect the presence of humans, while mapping in real-time the locations of the gases detected through a GPS module. A mobile robot combined with MOX gas sensors to detect early gas leaks in hazardous events was also developed by Palacín et al. (2019). The innovative system has 16 E-nose gas sensors that are low-cost. They can detect two chemicals at low concentrations that are located at broad distances from the source. In case two gas sources are present simultaneously, the mobile robot will detect the chemical that has the highest concentration. To detect and identify possible threats before damage takes place, Chen et al. (2023) proposed an AI-based monitoring system using processing parts that were designed to be simple. It had a 96.1% efficiency of detecting and localizing in real time external vibrational disturbance of a buried pipeline. Table 1 summarizes some important research studies that have used AI applications for monitoring and detection of hazardous substances.

Monitoring of hazardous environments to increase safety and to protect humans while they are operating in them was the goal of the work of Cheung et al. (2018). These authors developed a wireless sensor network (WSN) monitoring system with sensor nodes that can gather information regarding hazardous gases, temperature and humidity. They are embedded into a building information model (BIM) that displays the safety status in real-time with colors and shows the precise location of a possible dangerous event in advance. When high concentrations of gases are detected, the system is activated automatically using a flashing alarm that can warn the staff and workers. A ventilator and other safety devices are enabled to reduce toxic and flammable gas flow, which will prevent hazardous accidents. In a similar way, Jualayba et al. (2018) designed a monitoring and warning system that can display the safety status with colors when different levels of gases are detected. An exhaust fan is triggered when a medium level is displayed. At dangerous levels, an alarm buzzer is activated to inform people of a gas leakage and the need to decrease the concentration of the detected gas. The system has sensors for hydrogen, liquefied petroleum gas, and methane.

Wilson (2012) suggested a multifunctional gas detection system utilizing an Arduino Uno R3 and an Internet of Things module equipped with an MQ-6 gas sensor, capable of identifying methane, propane, and butane. The primary objective is to avert potential dangers in industrial settings. Upon surpassing predetermined gas threshold levels, the system activates a combination of a light-emitting diode (LED), buzzer, and notification message. There have been additional AI-powered devices devised for identifying hazardous gases in factories and industrial environments, all geared towards enhancing employee safety. They can collect data in real-time and identify the source of the gas leak and map the location (Manes et al., 2016). They are built with climatic sensors in addition to the gas sensors. They offer accurate, continuous monitoring of gas concentrations and are cost-effective (Thomas et al., 2018). Wearable smart sensing devices attached to workers cloths can be used for continuous gas monitoring in different industrial and other toxic areas. They provide accurate and real-time data acquisition that allows users to take rapid safety measures (Antolín et al., 2017). For instance, Binajjaj et al. (2018) proposed a wearable gas sensor network provided with an MQ-7 sensor that can detect both CO and CH_4, accurately and rapidly, using an accelerometer and an on-demand on/off switch that reduces energy consumption. It has a GPS and communication module that can transmit messages to other employees.

Wu et al. (2011) developed "WE-Safe", a wearable IoT sensor node capable of promptly warning workers in hazardous environments. This device is based on LoRa wireless technology. It has one microcontroller unit (MCU) and sensors for detection of CO, CO_2, as well as other sensors for environmental conditions (temperature, humidity, and UV light). It gives real time data and has a low energy exhaustion. Other researchers have designed devices with accurate and real time detection using MQ2, MQ5 gas sensors, which are capable of detecting hazardous gases in people's houses. The devices alarm the users through a buzzer and a short message service (SMS) notification is sent to tell them that the gas leakage needs to be controlled (Karthika et al., 2019; Panganiban, 2019).

Another application based on AI technology, as a strategy to control indoor ventilation by rapidly getting rid of PM, was developed by Kim et al. (2021). The ventilation system has a high removal rate of the hazardous airborne particles, which is done in a short time with low power consumption. Mendil et al. (2022) proposed a machine learning (ML)-based surrogate model for transport and dispersion of air pollutants that can predict, fast and accurately, the concentration and dose of pollutants in urban areas. Asha et al. (2022) designed an IoT and ML based model (ETAPM-AIT) for air quality monitoring that uses a sensor array for eight pollutants, such as NH_3, CO, NO_2, CH_4, CO_2, and $PM_{2.5}$. It also measures temperature and humidity. In this model an AAA-

TABLE 1 Research studies for AI applications in hazardous environments.

Detected parameter	Technology/Device used	Advantages	References
CO, CH_4	Wearable gas sensor network, Wireless sensor network, MQ-7 sensor, Arduino module with Xbee module	Accurate measurement of toxic gases, in time warning messages, low cost, low maintenance, reduced power consumption	Binajjaj et al. (2018)
Ethanol, acetone	Assistant Personal Robot (APR-02), 16 MOX gas sensors (e-nose), PLS-DA classifier	Autonomus, low cost, real-time, early detection, low gas concentration detection, two chemicals detection	Palacín et al. (2019)
CO_2, liquefied petroleum gas (LPG), vaporized alcohol (ethanol) gas, ambient gas	Six wheeled rocker-bogie robot, MQ gas sensors, HC-SR501 passive infrared (PIR) motion detector, HCSR04 ultrasonic sensors, Zigbee	98% accuracy for hazardous gases recognition, real-time gas detection, remote handling, obstacles avoidance, GPS location, human detection	Das et al. (2020)
CO, NO_2	SmokeBot platform, Mobile Robotic Olfaction (MRO) system, MOX gas sensors (UWAR nose)	Gas distribution mapping and navigation in environments with low visibility, accuracy of gas discrimination, search and rescue first responders' protection, less computational power	Fan et al. (2019)
Chlorofluorocarbon, CO, CH_4	Robotic vehicle, MQ sensors, Arduino UNO (R3)	Autonomous, obstacle avoidance, gas detection and degasification	Joshna et al. (2019)
C_2H_5OH, CO, H_2	Quad-rotor Unmanned Aerial Vehicle UAV, STM32F1 control chip, MQ-2 gas sensor, communication module NRF24L01 and SPI protocol	Autonomous navigation in dangerous environments, real-time detection of hazardous gases, wireless transmission	Shi et al. (2016)
Hydrogen, methane, Liquefied Petroleum Gas (LPG)	Hazardous gases detection and notification system, MQ gas sensors, Arduino, GSM module, indicator lamps	Gas leak SMS notifications, LCD display of detected gas level, buzzer alarm, exhaust fan to lower the gas concentration	JUALAYBA et al. (2018)
VOC, H_2S	Gas monitoring platform, gas and weather-climatic sensors, wireless sensor network (WSN), LAN/Ethernet (IEEE 802.1) with TCP/IP protocols, ARM Cortex-M3 32-bit micro-controller	Cost-effective, real-time emission identification, redeployable monitoring stations, continuous gas monitoring, rapid warning of hazardous events	Manes et al. (2016)
Particulate matter, CO, O_3 NO_2, noise, temperature, humidity	Monitor sensor network, Dragino WiFi IoT module HE, microprocessor (ATmega32u4)	Continuous monitoring of multiple hazards, low cost, accurate measurements	Thomas et al. (2018)
CO_2, CO, ultraviolet (UV), temperature, humidity	WE-Safe Platform, IoT sensor node, LoRa wireless technology, Arduino Uno, ATmega328p as MCU	Real-time data acquisition, remote cloud server, early warnings for workers in dangerous environments, low power consumption	Wu et al. (2011)
LPG, CO	Smart gas leakage, IoT with Arduino mega 2560, MQ-2 gas sensor, "BLYNK" mobile app, Wi-Fi module, Node-MCU ESP-8266	Gas leakage notification, gas level monitoring, automatic safety system, cost-effective	Zinnuraain et al. (2019)
H_2, LPG, CH_4, CO, alcohol	Gas detection system, Arduino UNO, MQ-5 Sensor	High sensitivity, fast response time, continuous update, alarm buzzer, LCD display, low cost	Bazrafshan and Kord Mostafapoor (2011)
NH_3, CO, NO_2, CH_4, CO_2, $PM_{2.5}$, temperature, humidity	ETAPM-AIT model based on IoT, ML techniques, WiFi, Artificial Algae Algorithm (AAA), Elman Neural Network (ENN) model	Real-time monitoring, alarm when hazardous substances are above limit, sensor array that can detect 8 parameters, cost-effective	Asha et al. (2022)
Temperature, ph, turbidity, conductivity, dissolved oxygen	IoT, Raspberry PI B+ core controller, WiFi, cloud computing, water monitoring sensors	Real-time monitoring, efficient processing, low cost	Vijayakumar and Ramya (2015)
pH, turbidity, temperature, humidity	Water quality monitoring system, IoT, ESP8266 Wi-Fi module, ThingSpeak mobile application, Arduino Mega	Efficient, real-time water quality monitoring, low cost	Pasika and Gandla (2020)
pH, temperature, conductivity	Water quality measurement machine, IoT, ESP8266 Wi-Fi module, ThingSpeak, Arduino Mega 2560 microcontroller	Real-time monitoring, highly accurate, track the level of water contamination, immediate warnings, low cost	Sarkar and Pandey (2015)
pH, dissolved oxygen, temperature, turbidity, conductivity	unmanned surface vehicle (USV) for water status monitoring, IoT, Wireless Sensor Network (WSN), ZigBee, IEEE 802.15.4 transceiver	Real-time data acquisition, remote handling in unreachable areas, high accuracy, efficient, reduced power consumption, cost-effective	Vasudevan and Baskaran (2021)

(Continued on following page)

TABLE 1 (*Continued*) Research studies for AI applications in hazardous environments.

Detected parameter	Technology/Device used	Advantages	References
Temperature, pH, turbidity, conductivity, DO, TDS, ORP	Water Quality Monitoring System, IoT, Arduino Nano, ESP8266 Wi-Fi module, ANN and SVM machine learning algorithms, ThingSpeak database	Remote monitoring, real-time data gathering, accurate water monitoring, automated water treatment corrective measure	Agrawal et al. (2017)
pH, temperature, turbidity, dissolved oxygen, conductivity	Smart Water Monitoring System (SWMS), four sensors, Arduino UNO, Raspberry Pi (RPi)	Real-time remote monitoring, SMS notification and alert signal, prevention of health hazards	Jha et al. (2018)

based ENN is used to classify air pollutants, and the model provides real-time detection of the pollutants. It has an alarm option that can alert people when hazardous pollutants are detected in the air. For safety purposes, Seo and Lee (2022) used an algorithm based on a CNN to accurately classify hazardous compounds to avoid hazardous chemicals being improperly used in laboratory environments. An intelligent system equipped with different sensors for CO, CO_2, NO_2, CH_4, and H_2S measurements using an Arduino Uno card and IoTs module monitors landfill biogas to prevent potential hazardous emissions (Mabrouki et al., 2021). Oduah and Ogunye (2023) developed a low-cost, smart, remote-sensing septic tank that can be used onsite to prevent sewage overflow. It uses an ultrasonic sensor for detection and monitoring the wastewater level in the septic tank and a GSM module to send SMS alerts to the users, avoiding spills of contaminants that can cause health problems.

Seker (2022) developed a real-time smart, cost-effective waste collection system that employs IoT with technologies like radio frequency identification (RFID), GIS, and ground penetrating radar systems (GPRS). It is effective for municipal waste collection and transportation and has the goal of reducing environmental pollution. Transportation and storage of hazardous materials can cause considerable risk to human life and the environment, due to their toxicity, corrosiveness, explosiveness, and radioactive characteristics (Mabrouk et al., 2017; Paredes-Belmar et al., 2017). In this respect, to design of a remote monitoring system based on a wireless sensor network that can offer real-time information on the vehicle that is transporting the hazardous materials. It can report the status of the vehicle (its attitude and tire pressure) and its location. In addition, the temperature of the transported hazardous chemicals is reported. This remote monitoring of hazardous materials during transportation provides safety measures for the workers and for citizens, minimizing the occurrence of accidents.

Hosseini and Verma (2021) proposed an analytical method that can determine the best route for transportation of hazardous materials, to assure the safety of inhabitants. It can exclude route zones that have high population density. Over the years, pollution of water resources has increased, making uncontaminated water bodies even more scarce all over the globe. The pollution has destroyed aquatic ecosystems and resulted in health and safety issues for humanity (Zhai et al., 2021; Yusuf et al., 2023). Protecting water resources and preventing the release of different pollutants represent the first steps in ensuring safety for the environment and people (Berman et al., 2020). Various approaches for monitoring the physico-chemical parameters of water resources have led to the implementation of novel smart technologies (Geetha and Gouthami, 2016; Huang et al., 2018). Water-quality smart monitoring systems using the Internet of Things (IoT) has gained popularity by engaging WSNs that offer continuous, remote and real-time monitoring and measurement of different water parameters. They can identify pollution sources and give the opportunity of early warnings, when possible threats or hazards are detected, and they can provide proper safety measures (Pule et al., 2017; Lakshmikantha et al., 2021).

A low cost system making use of the IoT for detection of water parameters in real time, like temperature, turbidity, conductivity, pH, and dissolved oxygen, which uses raspberry PI B+ as a core controller, has been designed (Vijayakumar and Ramya, 2015). Subsequently, Pasika and Gandla (Pasika and Gandla, 2020) proposed a low cost, efficient system based on the IoT using ThinkSpeak and various sensors to monitor drinking water quality in real time. In a similar way, Kumer et al. (Kumer et al., 2021) used ThinkSpeak and an Arduino Mega 2560 microcontroller in a device to analyze in real-time water parameters to detect levels of contamination when it happens and to warn inhabitants of hazardous health risks. It is very accurate and inexpensive.

Khan et al. (2020) developed a system that can monitor water quality for an industrial effluent treatment plant. It is composed of wireless sensor networks, a GSM module for notifications in case of emergency, an Arduino Uno R3 microcontroller, and an IoT-based cloud server. This system proved to be efficient and cost-effective. It gives continuous, real-time water quality monitoring, which allows the local authorities the benefit of supervising and checking to see if the released water from a specific industry is polluted or not. To tackle the problem of unreachable water bodies that are in secluded areas, Vasudevan and Baskaran (2021) proposed an innovative unmanned surface vehicle (USV) based on the IoT and sensors that provided control of these waters and diminished water pollution. They confirmed its efficiency, low investment, reduced energy consumption, and real-time data acquisition for surface water quality monitoring. Similarly, Ryu (2022) proposed an unmanned aircraft system (UAS) for water sampling and monitoring, which can be used in potentially hazardous environments, minimizing the involvement of humans. Its UASWQP platform can take water samples from various points and is effective in presenting real-time data regarding measurements of water parameters.

Adeleke et al. (2023) developed a highly accurate water quality monitoring system interfaced with sensors and machine learning algorithms to predict the level of water pollution. The remote system gathers water quality data in real-time and is built with an automated control system that can apply water treatment when the level of pollutants are above the standard limits, which reduces the expansion of diseases from contaminated water (Adeleke et al.,

TABLE 2 AI models for monitoring the environment for hazardous substances.

Environment	Approach/Models	Utility	Authors
Air	Hybrid Hidden Markov Model (HMM) - ANN	Harmful gas monitoring and error detection in sensor datasets	Praveenchandar et al. (2022)
Air	Artificial Algae Algorithm (AAA) based Elman Neural Network (ENN)	IoT enabled environmental toxicology for air pollution monitoring; classification of air pollutants	Asha et al. (2022)
Air	Random Forest (RF), Bagged Classification and Regression Trees (Bagged CART), and Mixture Discriminate Analysis (MDA)	Hazard prediction of particulate matter (PM_{10})	Choubin et al. (2020)
Air	Spatiotemporal deep learning (STDL)-based method using Stacked autoencoder (SAE) model	Air quality predictions	Li et al. (2016)
Soil	ANN, adaptive neuro-fuzzy inference system (ANFIS) models, multiple linear regression (MLR)	Lead and cadmium estimation from clay, organic carbon, pH, phosphorus, and total nitrogen	Bazoobandi et al. (2022)
Soil	ANN	Determine essential heavy metals based on Ca, K, and Mg concentrations	Sari et al. (2022)
Soil	Hybrid model integrating least absolute shrinkage and selection operator (LASSO), genetic algorithm (GA) and error back propagation neural network (BPNN) with remote sensing imageries	Spatial distribution of heavy metals in soil	Shi et al. (2022)
Soil	ANN and Support Vector Regression (SVR)	Rapid estimation of soil trace/heavy metals and related decision-making	Tawabini et al. (2022)
Soil	SVM with mid-infrared (MIR) laser spectroscopy	*In Situ* detection of petroleum in soils	Galán-Freyle et al. (2020)
Water	Artificial neural networks (ANN)	Lead removal from aqueous solution	Yetilmezsoy and Demirel (2008)
Water	ANN	Describe adsorption of benzene, toluene, ethyl benzene, and xylene on iron nano particles for their effective removal from aqueous systems	Mahmoud et al. (2018)
Water	ANN, Support vector machine (SVM), adaptive neuro-fuzzy inference system (ANFIS), genetic algorithms (GA)	Modelling and optimization of electrochemical processes for water and wastewater treatment	Shirkoohi et al. (2022)
Water	Logistic regression (LR), random forest classifier, and K-nearest neighbours (KNN)	Monitoring system for onsite septic systems failure	Ravi and Johnson (2021)

2023). A smart water monitoring system that collects real-time data for water quality and quantity was designed by Jha et al. (2018). The system uses sensors connected to an Arduino UNOTM and a Raspberry PiTM(RPi) microprocessor. It can avoid hazardous events caused by contaminated water seepage into drinking water sources by sending SMS and e-mail notifications when the monitored water is contaminated, which gives the opportunity for authorities and consumers to take rapid safety measures.

5 Advancements and challenges

5.1 BeeTox AI and other models

In a recent study (Moreira-Filho et al., 2021), a novel web application called BeeToxAI was introduced. It utilizes AI to assess the acute toxicity of chemicals to *Apis mellifera*, commonly known as the honey bee. BeeToxAI offers users a comprehensive set of features, including the classification as toxic or non-toxic of acute contact toxicity and acute oral toxicity endpoints. In addition, confidence scores for the prediction are given along with a visual representation of the results through maps. These color-coded maps illustrate the relative contribution of chemical fragments to toxicity. Since pesticides contain toxic substances that are injurious to health (Demirel and Kumral, 2021), AI algorithms are needed for controlling and tracking them, to reduce their toxicity. Because nitrogen based chemical fertilizers increase the amount of nitrate in groundwater, AI might be used to control excessive nitrate in soil.

The analysis of water quality as a source of irrigation can be done using AI-ANN models (Ostad-Ali-Askari et al., 2017). Over the past decade, a multitude of studies have been conducted to forecast the efficacy of machine learning in removing heavy metals from soil (Table 2). These investigations have focused on predicting the effectiveness of various techniques for soil remediation and heavy metal removal utilizing machine learning methodologies (Zafar et al., 2017; Zhu et al., 2019; Bhagat et al., 2020). Models used for optimizing and predicting heavy metal removal have encompassed black box, fuzzy logic, kernel, evolutionary, and hybrid models, and each offers distinct advantages.

In a study conducted by Talebkeikhah et al. (2022), which focussed on the adsorption of Pb (II) on biochar, eleven models were employed, which included group-data handling methods such as support vector machines (SVM), radiofrequency (RF), adaptive neuro-fuzzy interference systems (ANFIS), multilayer perception (MLP), and Decision Tree. Additionally, to evaluate the significance of coupling in creating predictive models for estimating adsorption

efficiency, the researchers developed four coupled predictive models using the grasshopper optimization algorithm (GOA) and the bat algorithm, and they combined ANFIS and MLP.

Chen et al. (2020) developed a CNN architecture specifically designed for deep calibration using near-infrared (NIR) spectroscopy data. Their study focused on assessing the level of water pollution originating from domestic and industrial sources, with the aim of enabling suitable agricultural irrigation practices. The researchers successfully established intelligent spectroscopic models using the CNN architecture, which could be instrumental in addressing water recycling and conservation issues in agricultural cultivation.

The field of medicine is experiencing gradual transformations due to the development of AI. Various medical disciplines, including clinical, diagnostic, rehabilitative, surgical, and predictive practices, are being influenced by AI applications (Secinaro et al., 2021). AI technologies have the capacity to process and analyze large volumes of data from diverse modalities to aid disease detection and guide clinical decision-making (Cho et al., 2020).

5.2 Medical waste

Recently, global healthcare advancements have led to an increase in the generation of medical waste, thereby contributing to a rising trend in waste production (Bazrafshan and Kord Mostafapoor, 2011). Various types of hazardous waste, such as hospital waste, dental waste from medical laboratories, blood wastage, and clinical waste, pose potential risks to both human health and the environment (Bazrafshan and Kord Mostafapoor, 2011). Medical waste (MW) encompasses infectious waste, sharp waste, toxic waste, chemical waste, and pharmaceutical waste. Inadequate management of medical waste can lead to the transmission of infectious diseases, including AIDS, hepatitis, typhoid, and many others (Askarian et al., 2010; Aghapour et al., 2013). These wastes have the potential to pollute the environment and spread acute and latent viral infections. They contain significant amounts of bacteria and viruses. If they are not properly controlled, they have the potential to contaminate the environment, which includes air, water, plants, animals, and land (Rajan et al., 2019). This can lead to the spread of disease. The quality of life, physical and emotional health, and overall health of medical personnel and patients are all seriously threatened by MW (Nwachukwu et al., 2013).

In a recent study (Zhou et al., 2022), a novel image recognition system was introduced called Deep MW for the purpose of sorting medical waste. Deep MW utilizes a CNN as its underlying architecture. The system aims to enhance the ease, accuracy, and efficiency of medical waste sorting and recycling processes, while also reducing the risk of occupational exposure for workers in medical waste facilities. Other similar systems exist, like iWaste (Chen et al., 2020), Deep MW excels in recognizing and sorting diverse categories of medical waste, showcasing remarkable accuracy and practicality in classification. Its proficiency makes it adaptable for expanding its use to other object classifications. Amidst the ongoing COVID-19 pandemic and other outbreaks, artificial intelligence has proven pivotal in creating surveillance tools aimed at identifying individuals not adhering to quarantine regulations. Monitoring bracelets exemplify one such tool. Furthermore, AI-enhanced technologies, such as smartphones and thermal cameras, are utilized for detecting fever and signs of illness. (Whitelaw et al., 2020).

5.3 Other applications

These advancements in AI have significantly contributed to improving surveillance and monitoring capabilities in public health and safety. A country like Taiwan treats coronavirus patients based on their travel history and symptoms by integrating data from the immigration and customs database with the national medical insurance database (Wang J. et al., 2023). Based on research conducted by Taiwan's Occupational Safety and Health Administration (OSHA), a significant number of surgical staff, approximately 50,000 individuals, face exposure to smoke hazards during procedures involving the use of lasers and electrocautery (Wang et al., 2020). This exposure poses potential risks to the health and safety of surgical personnel. AI can be used for assessing the characterization of the smoke emitted and also for carrying out an AI-powered air quality assessment (Kaginalkar et al., 2021).

Industry 4.0 is revolutionizing the way companies manufacture, improve, and distribute their products, and AI is widely recognized as a crucial technology for it (Mao et al., 2019). Its incorporation in these domains has significantly transformed and enhanced various aspects of industrial processes, enabling advanced automation, data-driven decision-making, predictive maintenance, and optimization of manufacturing operations (Figure 3). AI's capabilities have played a pivotal role in driving efficiency, productivity, and innovation in the manufacturing sector, making it a key enabler of the Industry 4.0 paradigm. Industry 4.0 has brought about a number of cutting-edge technologies, and this new phase of AI development is known as AI 2.0 (Pan, 2016). The utilization of AI in smart factories and its application in modern industrial sectors have been acknowledged as transformative (Mao et al., 2019). It enables efficient decision-making processes based on real-time and historical data, minimizing the need for human intervention.

Liu et al. (2019) proposed the integration of intelligent manufacturing in-shop service modules with Big Data analytics, to enhance productivity in various industries. For the facial mask industry, Zhang (2019) suggested a remote design system utilizing AI for image processing and pattern recognition, coupled with a client wireless communication. Yan et al. (2018) highlighted the significance of AI algorithms in forecasting the lifespan of industrial equipment, particularly in the context of Industry 4.0.

AI, machine learning, and autonomous technologies have significantly impacted the mining industry by detecting hazards and reducing risks [136, 137]. The introduction of autonomous trucks and the utilization of AI and machine learning have brought cost savings, increased productivity, improved worker safety, and continuous production (Hyder et al., 2018). AI systems analyzing geological, topographical, mineralogical, and mapping data can identify anomalies, locate potential areas of interest, and facilitate autonomous drilling operations (Hyder et al., 2018).

In the architecture, engineering, and construction (AEC) sector, AI has been increasingly employed to enhance existing procedures and address challenges. Applications and algorithms of AI in the AEC industry have been studied. Virtual reality and augmented reality have been used for hazard identification and risk assessment

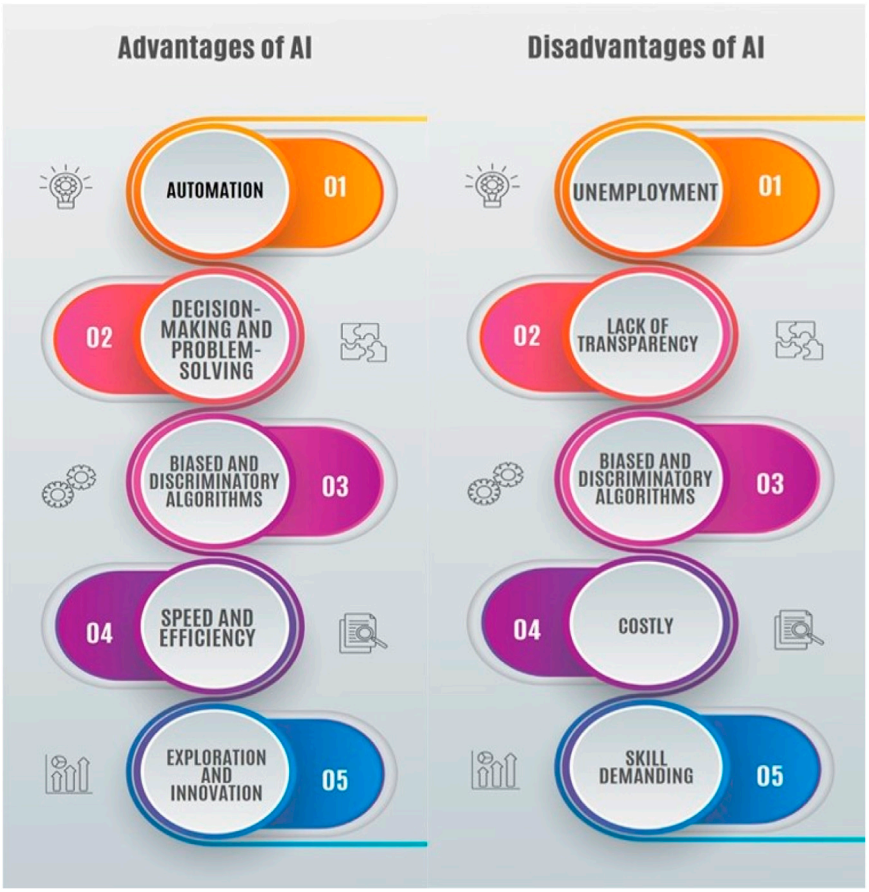


FIGURE 3
The figure illustrates the advantages and disadvantages associated with AI.

in construction environments (Von Meding et al., 2009; Perlman et al., 2014; Abioye et al., 2021).

Advancements in artificial intelligence (AI) and machine learning (ML) have significantly improved efforts in mitigating pollution (Chen et al., 2023). State-of-the-art technologies are being utilized to improve air quality modeling, integrating a variety of data sources such as satellite imagery and meteorological data for more accurate predictions. The amalgamation of AI with satellite technology allows for comprehensive monitoring of environmental shifts, assisting in the identification of pollution sources. The introduction of edge computing enables real-time analysis of environmental data at its origin, facilitating swift responses to pollution incidents. Hybrid models, which merge physics-based simulations with ML, enhance the precision of identifying pollution sources (Anthony et al., 2023). ML-driven predictive analytics forecast pollution levels, enabling proactive measures to prevent environmental deterioration. In the energy sector, AI is optimizing smart grids and energy distribution, reducing the environmental impact (Gautam et al., 2023). AI-guided robotics are deployed for pollution clean-up, autonomously recognizing and addressing contaminants. The combination of AI and block chain technology ensures transparency and traceability in environmental data (Zhao et al., 2023). ML-powered adaptive monitoring networks intelligently allocate resources for effective pollution monitoring. Mobile applications, leveraging AI, empower citizens to contribute to environmental monitoring, promoting a decentralized approach to data collection (Sahil et al., 2023). These recent strides underscore the transformative potential of AI and ML in reshaping pollution mitigation strategies, steering towards a more sustainable and resilient future (Zahed et al., 2022).

6 Limitations

In recent years, there has been a notable surge in the utilization of AI as an emerging technology across various sectors, such as agriculture, medicine and healthcare, mining and manufacturing, environmental conservation, and numerous other domains (Trasande et al., 2011; Perez Santin et al., 2021). While AI brings forth numerous benefits, it is not extempted from challenges and limitations that require efficient and notable solutions. These issues encompass the effective execution of tasks and management and mitigation of risks, as well as the assessment of hazardous materials. The requirement for significant processing power raises concerns about environmental impact and energy consumption (Singh and Kaur, 2022). Developing countries may face infrastructure limitations that hinder the efficient deployment of AI (Singh and Kaur, 2022). Data acquisition costs in data-driven agriculture remain high, limiting the widespread impact of AI on

agricultural productivity (Linaza et al., 2021). Furthermore, issues of data ownership, privacy, and cybersecurity arise, necessitating standardized protocols and integration with technologies like block chain to ensure data integrity and security (Abedjan et al., 2019; Misra et al., 2020). Block chain is a system in which a record of transactions is maintained across computers that are linked.

In the manufacturing field, there is a need for further research on the application of AI in various sub-components, such as design, machinery used, assembly, and production (Zeba et al., 2021). Ethical considerations also arise, including concerns about genetic engineering, moral decision-making by AI systems, and the impact on jobs, economy, and society (Zeba et al., 2021), including privacy under surveillance. The AEC industry faces challenges related to funding, data, ethics, privacy, trust, scalability, and expertise (Motawa, 2017; Hagendorff, 2020). Resistance to technology adoption stems from concerns about job displacement, uncertainty, wealth distribution, and the complex relationship between humans and technology (Kappal, 2017). Additionally, there are challenges in understanding how AI can effectively support individual and group decision-making (Maedche et al., 2019). Overall, the application of AI in various industries brings about significant opportunities and challenges, necessitating further research, infrastructure development, and ethical considerations.

7 Conclusion and future perspectives

The development of artificial intelligence-driven sensors for environmental monitoring of hazardous substances has the potential to revolutionize how we detect and respond to environmental threats. With their ability to process large amounts of data in real-time and identify patterns and anomalies that may indicate the presence of hazardous substances, these sensors can greatly improve our ability to protect public health and the environment.

However, there are still many challenges that must be addressed, such as ensuring the accuracy and reliability of these sensors and determining the best ways to integrate them into existing monitoring systems. Despite these challenges, the potential benefits of AI-driven sensors make them a promising avenue for future research and development in the field of environmental monitoring. In the near future, AI-powered hazardous substance monitoring has the potential for enhanced automation, the creation of more sophisticated models, and the fusion of AI with novel technologies like the IoT.

AI could deliver promises about prediction, optimization, and decision-making in the near future to help the traditional construction industry keep up with the rapid speed of automation and digitalization in a variety of risky and dangerous circumstances (Pan, 2016). It can offer a wide range of risk identification factors as well as assess and deliver for prioritization, because it can monitor, recognize, analyze, and anticipate possible risk in terms of safety, quality, efficiency, and cost across teams and work areas even in the presence of significant uncertainty (Afzal et al., 2021). AI-powered monitoring systems can provide real-time, on-ground detection of potentially dangerous substances and the location of the incident sites. They may prevent any serious accidents from happening that would be harmful to human health. These systems are applicable in the fields of agriculture, health, mining, manufacturing, and other industries (Ghayvat et al., 2021; Kim et al., 2021).

AI has the potential to remodel the monitoring and detection of hazardous substances in various environments by offering real-time analysis. This can be achieved through the integration of sensors, data processing capabilities, and machine learning algorithms. By leveraging these technologies, AI systems can swiftly and accurately analyse data streams to identify potential hazards. This application is particularly valuable in industrial settings or areas susceptible to chemical spills or leaks (Leppert et al., 2012). Furthermore, AI can contribute to predictive modelling and risk assessment of hazardous substances. By examining historical data and considering various environmental factors, AI algorithms can forecast the probability and severity of incidents involving hazardous substances. This allows for proactive mitigation measures to be implemented. AI can also integrate data from multiple sources, such as sensors, satellite imagery, and public health records. By combining diverse datasets, AI algorithms provide a holistic understanding of the distribution, movement, and impacts of hazardous substances (Wu et al., 2011; Hurlbert et al., 2019; Shafique et al., 2022; Shi et al., 2022). In addition, AI can offer decision support tools for human operators engaged in hazardous substance monitoring. Through user-friendly interfaces and natural language processing, AI systems can aid in data interpretation, anomaly detection, and timely decision-making. These features assist operators in effectively managing hazardous substance situations.

Author contributions

SP: Investigation, Visualization, Writing–original draft. SM: Conceptualization, Visualization, Writing–original draft, Writing–review and editing. OW: Writing–original draft, Writing–review and editing. SK: Writing–original draft. VS: Investigation, Visualization, Writing–original draft. AS: Writing–original draft. VA: Writing–original draft. MK: Data curation, Formal Analysis, Investigation, Writing–original draft, Writing–review and editing. DH: Formal Analysis, Validation, Visualization, Writing–original draft, Writing–review and editing. NB: Formal Analysis, Investigation, Supervision, Visualization, Writing–original draft, Writing–review and editing. YC: Formal Analysis, Funding acquisition, Investigation, Resources, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The Basic Science Research Program supported this research through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2019R1A6A1A11052070).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abedjan, Z., Boujemaa, N., Campbell, S., Casla, P., Chatterjea, S., Consoli, S., et al. (2019). *Data science in healthcare: benefits, challenges and opportunities*. Springer.

Abioye, S. O., Oyedele, L. O., Akanbi, L., Ajayi, A., Delgado, J. M. D., Bilal, M., et al. (2021). Artificial intelligence in the construction industry: a review of present status, opportunities and future challenges. *J. Build. Eng.* 44, 103299. doi:10.1016/j.jobe.2021.103299

Adeleke, I. A., Nwulu, N. I., and Ogbolumani, O. A. (2023). A hybrid machine learning and embedded IoT-based water quality monitoring system. *Internet Things* 22, 100774. doi:10.1016/j.iot.2023.100774

Afzal, F., Yunfei, S., Nazir, M., and Bhatti, S. M. (2021). A review of artificial intelligence based risk assessment methods for capturing complexity-risk interdependencies: cost overrun in construction projects. *Int. J. Manag. Proj. Bus.* 14 (2), 300–328. doi:10.1108/ijmpb-02-2019-0047

Aggestam, F., and Mangalagiu, D. (2020). Is sharing truly caring? Environmental data value chains and policymaking in Europe and Central Asia. *Environ. Sci. Policy* 114, 152–161. doi:10.1016/j.envsci.2020.07.012

Aghapour, P., Nabizadeh, R., Nouri, J., Monavari, M., and Yaghmaeian, K. (2013). Analysis of the health and environmental status of sterilizers in hospital waste management: a case study of hospitals in Tehran. *Waste Manag. Res.* 31 (3), 333–337. doi:10.1177/0734242x12472706

Agrawal, A., Kumar, L., Kumar, P., and Jha, V. K. (2017). IOT based hazardous gas detection system using AVR microcontroller. *Int. Res. J. Eng. Technol.* 4 (3), 1–5.

Ahmadi, M., Jorfi, S., Birgani, Y. T., and Ghaedrahmat, Z. (2021). Monitoring and application of artificial neural network model for prediction of organophosphorus pesticides residue in Ahvaz water treatment plants. *Biointerface Res. Appl. Chem.* 11 (6), 14032–14043.

Ahmed, S. A., Popov, V. L., Topalov, A. V., and Shakev, N. G. (2016). Perimeter detection and surveillance of polluted areas by robotized agents in a hybrid wireless sensor network. *IFAC-PapersOnLine* 49 (29), 247–252. doi:10.1016/j.ifacol.2016.11.061

Ai, W., Chen, G., Yue, X., and Wang, J. (2023). Application of hyperspectral and deep learning in farmland soil microplastic detection. *J. Hazard. Mater.* 445, 130568. doi:10.1016/j.jhazmat.2022.130568

Amado, T. M., and Cruz, J. C. D. (2018). "Development of machine learning-based predictive models for air quality monitoring and characterization," in TENCON 2018-2018 IEEE Region 10 Conference.

Anthony, D., Siriwardana, H., Ashvini, S., Pallewatta, S., Samarasekara, S. M., Edirisinghe, S., et al. (2023). Trends in marine pollution mitigation technologies: scientometric analysis of published literature (1990-2022). *Regional Stud. Mar. Sci.* 66, 103156. doi:10.1016/j.rsma.2023.103156

Antolín, D., Medrano, N., Calvo, B., and Pérez, F. (2017). A wearable wireless sensor network for indoor smart environment monitoring in safety applications. *Sensors* 17 (2), 365. doi:10.3390/s17020365

Asha, P., Natrayan, L., Geetha, B., Beulah, J. R., Sumathy, R., Varalakshmi, G., et al. (2022). IoT enabled environmental toxicology for air pollution monitoring using AI techniques. *Environ. Res.* 205, 112574. doi:10.1016/j.envres.2021.112574

Askarian, M., Heidarpoor, P., and Assadian, O. (2010). A total quality management approach to healthcare waste management in Namazi Hospital, Iran. *Waste Manag.* 30 (11), 2321–2326. doi:10.1016/j.wasman.2010.06.020

Bachmann, T. M. (2006). *Hazardous substances and human health: exposure, impact and external cost assessment at the European scale*. Elsevier.

Baran, W., Adamek, E., Ziemiańska, J., and Sobczak, A. (2011). Effects of the presence of sulfonamides in the environment and their influence on human health. *J. Hazard. Mater.* 196, 1–15. doi:10.1016/j.jhazmat.2011.08.082

Basel, A. M., Nguyen, K. T., Arnaud, E., and Craparo, A. C. (2023). The foundations of big data sharing: a CGIAR international research organization perspective. *Front. Environ. Sci.* 11, 1107393. doi:10.3389/fenvs.2023.1107393

Bazoobandi, A., Emamgholizadeh, S., and Ghorbani, H. (2022). Estimating the amount of cadmium and lead in the polluted soil using artificial intelligence models. *Eur. J. Environ. Civ. Eng.* 26 (3), 933–951. doi:10.1080/19648189.2019.1686429

Bazrafshan, E., and Kord Mostafapoor, F. (2011). Survey of medical waste characterization and management in Iran: a case study of Sistan and Baluchestan Province. *Waste Manag. Res.* 29 (4), 442–450. doi:10.1177/0734242x10374901

Benger, M. J. (2004). The EPA remote sensing archive," in *The Landscape Science Seminar Series*. Reston, VA: Video Teleconference.

Berman, I., Zereik, E., Kapitonov, A., Bonsignorio, F., Khassanov, A., Oripova, A., et al. (2020). Trustable environmental monitoring by means of sensors networks on swarming autonomous marine vessels and distributed ledger technology. *Front. Robotics AI* 7, 70. doi:10.3389/frobt.2020.00070

Bhagat, S. K., Tung, T. M., and Yaseen, Z. M. (2020). Development of artificial intelligence for modeling wastewater heavy metal removal: state of the art, application assessment and possible future research. *J. Clean. Prod.* 250, 119473. doi:10.1016/j.jclepro.2019.119473

Bharati, M. S. S., and Soma, V. R. (2021). Flexible SERS substrates for hazardous materials detection: recent advances. *Opto-Electronic Adv.* 4 (11), 210048. doi:10.29026/oea.2021.210048

Binajjaj, A., Sheltami, T., Aliyu, F., and Kaosar, M. (2018). Design and implementation of a wearable gas sensor network for oil and gas industry workers. *J. Comput.* 13 (3), 300–308. doi:10.17706/jcp.13.3.300-308

Bolan, N., Sarkar, B., Yan, Y., Li, Q., Wijesekara, H., Kannan, K., et al. (2021). Remediation of poly-and perfluoroalkyl substances (PFAS) contaminated soils–To mobilize or to immobilize or to degrade? *J. Hazard. Mater.* 401, 123892. doi:10.1016/j.jhazmat.2020.123892

Bølviken, B., Honey, F., Levine, S., Lyon, R., and Prelat, A. (1977). Detection of naturally heavy-metal-poisoned areas by Landsat-1 digital data. *Dev. Econ. Geol.* 9, 457–471. Elsevier.

Burgués, J., Hernández, V., Lilienthal, A. J., and Marco, S. (2019). Smelling nano aerial vehicle for gas source localization and mapping. *Sensors* 19 (3), 478. doi:10.3390/s19030478

Chen, H., Chen, A., Xu, L., Xie, H., Qiao, H., Lin, Q., et al. (2020). A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. *Agric. Water Manag.* 240, 106303. doi:10.1016/j.agwat.2020.106303

Chen, H., Wong, R. C.-K., Park, S., and Hugo, R. (2023). An AI-based monitoring system for external disturbance detection and classification near a buried pipeline. *Mech. Syst. Signal Process.* 196, 110346. doi:10.1016/j.ymssp.2023.110346

Cheung, W.-F., Lin, T.-H., and Lin, Y.-C. (2018). A real-time construction safety monitoring system for hazardous gas integrating wireless sensor network and building information modeling technologies. *Sensors* 18 (2), 436. doi:10.3390/s18020436

Cho, B.-J., Choi, Y. J., Lee, M.-J., Kim, J. H., Son, G.-H., Park, S.-H., et al. (2020). Classification of cervical neoplasms on colposcopic photography using deep learning. *Sci. Rep.* 10 (1), 13652–13710. doi:10.1038/s41598-020-70490-4

Choubin, B., Abdolshahnejad, M., Moradi, E., Querol, X., Mosavi, A., Shamshirband, S., et al. (2020). Spatial hazard assessment of the PM10 using machine learning models in Barcelona, Spain. *Sci. Total Environ.* 701, 134474. doi:10.1016/j.scitotenv.2019.134474

Cruz, F. R. G., Binag, M. G., Ga, M. R. G., and Uy, F. A. A. (2018). "Flood prediction using multi-layer artificial neural network in monitoring system with rain gauge, water level, soil moisture sensors," in TENCON 2018-2018 IEEE Region 10 Conference.

Crystal-Ornelas, R., Varadharajan, C., O'Ryan, D., Beilsmith, K., Bond-Lamberty, B., Boye, K., et al. (2022). Enabling FAIR data in Earth and environmental science with community-centric (meta) data reporting formats. *Sci. data* 9 (1), 700. doi:10.1038/s41597-022-01606-w

Daam, M. A., Chelinho, S., Niemeyer, J. C., Owojori, O. J., De Silva, P. M. C., Sousa, J. P., et al. (2019). Environmental risk assessment of pesticides in tropical terrestrial ecosystems: test procedures, current status and future perspectives. *Ecotoxicol. Environ. Saf.* 181, 534–547. doi:10.1016/j.ecoenv.2019.06.038

Das, T., Sut, D. J., Gupta, V., Gohain, L., Kakoty, P., and Kakoty, N. M. (2020). "A mobile robot for hazardous gas sensing," in 2020 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 02-04 July 2020 (IEEE).

Demirel, M., and Kumral, N. A. (2021). "Artificial intelligence in integrated pest management," in *Artificial intelligence and IoT-based technologies for sustainable farming and smart agriculture* (Hershey, Pennsylvania, United States: IGI Global), 289–313.

Emaminejad, N., and Akhavian, R. (2022). Trustworthy AI and robotics: implications for the AEC industry. *Automation Constr.* 139, 104298. doi:10.1016/j.autcon.2022.104298

Errico, A., Angelino, C. V., Cicala, L., Podobinski, D. P., Persechino, G., Ferrara, C., et al. (2014). SAR/multispectral image fusion for the detection of environmental hazards with a GIS. *Earth Resour. Environ. Remote Sensing/GIS Appl. V*. doi:10.1117/12.2066476

Fan, H., Hernandez Bennetts, V., Schaffernicht, E., and Lilienthal, A. J. (2019). Towards gas discrimination and mapping in emergency response scenarios using a mobile robot with an electronic nose. *Sensors* 19 (3), 685. doi:10.3390/s19030685

Farrokhi-Asl, H., Makui, A., Ghousi, R., and Rabbani, M. (2020). Developing a hazardous waste management system with consideration of health, safety, and environment. *Comput. Electr. Eng.* 82, 106553. doi:10.1016/j.compeleceng.2020.106553

Fung, A. G., Rajapakse, M. Y., McCartney, M. M., Falcon, A. K., Fabia, F. M., Kenyon, N. J., et al. (2019). Wearable environmental monitor to quantify personal ambient volatile organic compound exposures. *ACS sensors* 4 (5), 1358–1364. doi:10.1021/acssensors.9b00304

Galán-Freyle, N. J., Ospina-Castro, M. L., Medina-González, A. R., Villarreal-González, R., Hernández-Rivera, S. P., and Pacheco-Londoño, L. C. (2020). Artificial intelligence assisted mid-infrared laser spectroscopy *in situ* detection of petroleum in soils. *Appl. Sci.* 10 (4), 1319. doi:10.3390/app10041319

Gallego, V., Rossi, M., and Brunelli, D. (2015). "Unmanned aerial gas leakage localization and mapping using microdrones," in 2015 IEEE Sensors Applications Symposium (SAS), Zadar, Croatia, 13-15 April 2015 (IEEE).

Garofalo, D. (2003). *Aerial photointerpretation of hazardous waste sites: an overview*. Boca Raton, FL, USA: CRC Press LLC.

Gautam, K., Sharma, P., Dwivedi, S., Singh, A., Gaur, V. K., Varjani, S., et al. (2023). A review on control and abatement of soil pollution by heavy metals: emphasis on artificial intelligence in recovery of contaminated soil. *Environ. Res.* 225, 115592. doi:10.1016/j.envres.2023.115592

Geetha, S., and Gouthami, S. (2016). Internet of things enabled real time water quality monitoring system. *Smart Water* 2 (1), 1–19. doi:10.1186/s40713-017-0005-y

Gerhardt, N., Clothier, R., and Wild, G. (2014). "Investigating the practicality of hazardous material detection using unmanned aerial systems," in 2014 IEEE Metrology for Aerospace (MetroAeroSpace), Benevento, Italy, 29-30 May 2014 (IEEE).

Ghayvat, H., Awais, M., Gope, P., Pandya, S., and Majumdar, S, (2021). Recognizing suspect and predicting the spread of contagion based on mobile phone location data (counteract): a system of identifying covid-19 infectious and hazardous sites, detecting disease outbreaks based on the internet of things, edge computing, and artificial intelligence. *Sustain. Cities Soc.* 69, 102798. doi:10.1016/j.scs.2021.102798

Hagendorff, T. (2020). The ethics of AI ethics: an evaluation of guidelines. *Minds Mach.* 30 (1), 99–120. doi:10.1007/s11023-020-09517-8

Haldorai, A., Suriya, M., and Balakrishnan, M. (2024). An improved single short detection method for smart vision-based water garbage cleaning robot. *Cogn. Robot.* 4, 19–29. doi:10.1016/j.cogr.2023.11.002

Hosseini, S. D., and Verma, M. (2021). Equitable routing of rail hazardous materials shipments using CVaR methodology. *Comput. Operations Res.* 129, 105222. doi:10.1016/j.cor.2021.105222

Huang, C., Chen, Y., Zhang, S., and Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: a review. *Rev. Geophys.* 56 (2), 333–360. doi:10.1029/2018rg000598

Hurlbert, M., Krishnaswamy, J., Johnson, F. X., Rodríguez-Morales, J. E., and Zommers, Z. (2019). *Risk management and decision making in relation to sustainable development*.

Hyder, Z., Siau, K., and Nah, F. F.-H. (2018). *Use of artificial intelligence, machine learning, and autonomous technologies in the mining industry*.

Jeong, J., and Choi, J. (2022). Artificial intelligence-based toxicity prediction of environmental chemicals: future directions for chemical management applications. *Environ. Sci. Technol.* 56 (12), 7532–7543. doi:10.1021/acs.est.1c07413

Jha, M. K., Sah, R. K., Rashmitha, M., Sinha, R., Sujatha, B., and Suma, a. (2018). "Smart water monitoring system for real-time water quality and usage monitoring," in 2018 International Conference on Inventive Research in Computing Applications (ICIRCA).

Jia, X., Cao, Y., O Connor, D., Zhu, J., Tsang, D. C., Zou, B., et al. (2021). Mapping soil pollution by using drone image recognition and machine learning at an arsenic-contaminated agricultural field. *Environ. Pollut.* 270, 116281. doi:10.1016/j.envpol.2020.116281

Jiang, J., Lin, C., Han, G., Abu-Mahfouz, A. M., Shah, S. B. H., and Martínez-García, M. (2022). How AI-enabled SDN technologies improve the security and functionality of industrial IoT network: architectures, enabling technologies, and opportunities. *Digital Commun. Netw.* 9, 1351–1362. doi:10.1016/j.dcan.2022.07.001

Joshna, V., Kashyap, M., Ananya, V., and Manitha, P. (2019). "Fully autonomous robot to detect and degasify hazardous gas after flood disaster," in 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC).

Jualayba, M., Regio, K., Quiozon, H., and Destreza, A. (2018). "Hazardous gas detection and notification system," in 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM).

Kaginalkar, A., Kumar, S., Gargava, P., and Niyogi, D. (2021). Review of urban computing in air quality management as smart city service: an integrated IoT, AI, and cloud technology perspective. *Urban Clim.* 39, 100972. doi:10.1016/j.uclim.2021.100972

Kappal, S. (2017). *The biggest challenges in implementing AI. DZone*.

Karthika, S., Vanitha, U., Rejina, P., Supraja, A., Sampritha, R., and Srinithi, K. (2019). Toxic gas detection using low power controller with alert system. *Int. J. Innov. Technol. Explor. Eng.* 8, 1074–1079. doi:10.35940/ijitee.i7655.0881019

Kemper, T., and Sommer, S. (2004). *Use of airborne hyperspectral data to estimate residual heavy metal contamination and acidification potential in the Guadiamar floodplain Andalusia, Spain after the Aznacollar mining accident*. Maspalomas, Canary Islands, Spain: Remote sensing for environmental monitoring, GIS applications, and geology IV.

Khan, A., Gupta, S., and Gupta, S. K. (2020). Multi-hazard disaster studies: monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques. *Int. J. disaster risk Reduct.* 47, 101642. doi:10.1016/j.ijdrr.2020.101642

Kim, N. K., Kang, D. H., Lee, W., and Kang, H. W. (2021). Airflow pattern control using artificial intelligence for effective removal of indoor airborne hazardous materials. *Build. Environ.* 204, 108148. doi:10.1016/j.buildenv.2021.108148

Kim, Y., Evans, R. G., and Iversen, W. M. (2008). Remote sensing and control of an irrigation system using a distributed wireless sensor network. *IEEE Trans. Instrum. Meas.* 57 (7), 1379–1387. doi:10.1109/tim.2008.917198

Kostal, J., Brooks, B. W., Smith, C. A., and Devineni, G. (2022). O data, where art thou? Revolutionizing data sharing to advance our sustainability goals through smart chemical innovation. *Iscience* 25 (11), 105256. doi:10.1016/j.isci.2022.105256

Kumar, S. S., Mir, S. A., Wani, O. A., Babu, S., Yeasin, M., Bhat, M., et al. (2022). Land-use systems regulate carbon geochemistry in the temperate Himalayas, India. *J. Environ. Manag.* 320, 115811. doi:10.1016/j.jenvman.2022.115811

Kumer, S. A., Kanakaraja, P., Mounika, V., Abhishek, D., and Reddy, B. P. (2021). Environment water quality monitoring system. *Mater. Today Proc.* 46, 4137–4141. doi:10.1016/j.matpr.2021.02.674

Lakshmikantha, V., Hiriyannagowda, A., Manjunath, A., Patted, A., Basavaiah, J., and Anthony, A. A. (2021). IoT based smart water quality monitoring system. *Glob. Transitions Proc.* 2 (2), 181–186. doi:10.1016/j.gltp.2021.08.062

Leppert, J., Horner, G., Rietz, F., Ringer, J., Lammers, P. S., and Boeker, P. (2012). Near real time detection of hazardous airborne substances. *Talanta* 101, 440–446. doi:10.1016/j.talanta.2012.09.056

Levy, R., Remer, L., Kleidman, R., Mattoo, S., Ichoku, C., Kahn, R., et al. (2010). Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmos. Chem. Phys.* 10 (21), 10399–10420. doi:10.5194/acp-10-10399-2010

Li, X., Peng, L., Hu, Y., Shao, J., and Chi, T. (2016). Deep learning architecture for air quality predictions. *Environ. Sci. Pollut. Res.* 23, 22408–22417. doi:10.1007/s11356-016-7812-9

Li, Z., Yang, Y., Chen, X., He, Y., Bolan, N., Rinklebe, J., et al. (2022). A discussion of microplastics in soil and risks for ecosystems and food chains. *Chemosphere* 313, 137637. doi:10.1016/j.chemosphere.2022.137637

Linaza, M. T., Posada, J., Bund, J., Eisert, P., Quartulli, M., Döllner, J., et al. (2021). Data-driven artificial intelligence applications for sustainable precision agriculture. *Agronomy* 11 (6), 1227. doi:10.3390/agronomy11061227

Liu, C., Zhou, Y., Cen, Y., and Lin, D. (2019). Integrated application in intelligent production and logistics management: technical architectures concepts and business model analyses for the customised facial masks manufacturing. *Int. J. Comput. Integr. Manuf.* 32 (4-5), 522–532. doi:10.1080/0951192x.2019.1599434

Mabrouk, A., Boulmakoul, A., Karim, L., and Lbath, A. (2017). Safest and shortest itineraries for transporting hazardous materials using split points of Voronoï spatial diagrams based on spatial modeling of vulnerable zones. *Procedia Comput. Sci.* 109, 156–163. doi:10.1016/j.procs.2017.05.311

Mabrouki, J., Azrour, M., Fattah, G., Dhiba, D., and El Hajjaji, S. (2021). Intelligent monitoring system for biogas detection based on the Internet of Things: mohammedia, Morocco city landfill case. *Big Data Min. Anal.* 4 (1), 10–17. doi:10.26599/bdma.2020.9020017

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., et al. (2019). AI-based digital assistants: opportunities, threats, and research perspectives. *Bus. Inf. Syst. Eng.* 61, 535–544. doi:10.1007/s12599-019-00600-8

Mahmoud, A. S., Mostafa, M. K., and Abdel-Gawad, S. A. (2018). Artificial intelligence for the removal of benzene, toluene, ethyl benzene and xylene (BTEX) from aqueous solutions using iron nanoparticles. *Water Supply* 18 (5), 1650–1663. doi:10.2166/ws.2017.225

Manes, G., Collodi, G., Gelpi, L., Fusco, R., Ricci, G., Manes, A., et al. (2016). Realtime gas emission monitoring at hazardous sites using a distributed point-source sensing infrastructure. *Sensors* 16 (1), 121. doi:10.3390/s16010121

Manickavasagan, A., Jayas, D., White, N., and Jian, F. (2006). Thermal imaging of a stored grain silo to detect a hot spot. *Appl. Eng. Agric.* 22 (6), 891–897. doi:10.13031/2013.22243

Mansoor, S., Khan, N. F., Farooq, I., Kaur, N., Manhas, S., Raina, S., et al. (2022). "Phytoremediation at molecular level," in *Phytoremediation* (Elsevier), 65–90.

Mansoor, S., Kour, N., Manhas, S., Zahid, S., Wani, O. A., Sharma, V., et al. (2021). Biochar as a tool for effective management of drought and heavy metal toxicity. *Chemosphere* 271, 129458. doi:10.1016/j.chemosphere.2020.129458

Mao, S., Wang, B., Tang, Y., and Qian, F. (2019). Opportunities and challenges of artificial intelligence for green manufacturing in the process industry. *Engineering* 5 (6), 995–1002. doi:10.1016/j.eng.2019.08.013

Mendil, M., Leirens, S., Armand, P., and Duchenne, C. (2022). Hazardous atmospheric dispersion in urban areas: a Deep Learning approach for emergency pollution forecast. *Environ. Model. Softw.* 152, 105387. doi:10.1016/j.envsoft.2022.105387

Mertikas, S. P., Partsinevelos, P., Mavrocordatos, C., and Maximenko, N. A. (2021). "Environmental applications of remote sensing," in *Pollution assessment for sustainable practices in applied sciences and engineering* (Elsevier), 107–163.

Messinger, M., and Silman, M. (2016). Unmanned aerial vehicles for the assessment and monitoring of environmental contamination: an example from coal ash spills. *Environ. Pollut.* 218, 889–894. doi:10.1016/j.envpol.2016.08.019

Misra, N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., and Martynenko, A. (2020). IoT, big data, and artificial intelligence in agriculture and food industry. *IEEE Internet things J.* 9 (9), 6305–6324. doi:10.1109/jiot.2020.2998584

Moreira-Filho, J. T., Braga, R. C., Lemos, J. M., Alves, V. M., Borba, J. V., Costa, W. S., et al. (2021). BeetoxAI: an artificial intelligence-based web app to assess acute toxicity of chemicals to honey bees. *Artif. Intell. Life Sci.* 1, 100013. doi:10.1016/j.ailsci.2021.100013

Motawa, I. (2017). Spoken dialogue BIM systems–an application of big data in construction. *Facilities* 35 (13/14), 787–800. doi:10.1108/f-01-2016-0001

Naz, H., Sayyed, R., Khan, R. U., Naz, A., Wani, O. A., Maqsood, A., et al. (2023). Mesorhizobium improves chickpea growth under chromium stress and alleviates chromium contamination of soil. *J. Environ. Manag.* 338, 117779. doi:10.1016/j.jenvman.2023.117779

Nwachukwu, N. C., Orji, F. A., and Ugbogu, O. C. (2013). Health care waste management–public health benefits, and the need for effective environmental regulatory surveillance in federal Republic of Nigeria. *Curr. Top. public health* 2, 149–178. doi:10.5772/53196

Oduah, U. I., and Ogunye, E. B. (2023). A smart solution for preventing environmental pollution caused by overflowing onsite sewage septic tank. *Heliyon* 9 (4), e14925. doi:10.1016/j.heliyon.2023.e14925

Ostad-Ali-Askari, K., Shayannejad, M., and Ghorbanizadeh-Kharazi, H. (2017). Artificial neural network for modeling nitrate pollution of groundwater in marginal area of Zayandeh-rood River, Isfahan, Iran. *KSCE J. Civ. Eng.* 21, 134–140. doi:10.1007/s12205-016-0572-8

Palacín, J., Martínez, D., Clotet, E., Pallejà, T., Burgués, J., Fonollosa, J., et al. (2019). Application of an array of metal-oxide semiconductor gas sensors in an assistant personal robot for early gas leak detection. *Sensors* 19 (9), 1957. doi:10.3390/s19091957

Pan, Y. (2016). Heading toward artificial intelligence 2.0. *Engineering* 2 (4), 409–413. doi:10.1016/j.eng.2016.04.018

Panganiban, E. (2019). Automated hazardous gas detecting robot using wireless sensor networks with GSM-SMS alert and fire control system for households. *Int. J. Adv. Trends Comput. Sci. Eng.* 8 (3), 804–809. doi:10.30534/ijatcse/2019/72832019

Pant, P., Lal, R. M., Guttikunda, S. K., Russell, A. G., Nagpure, A. S., Ramaswami, A., et al. (2019). Monitoring particulate matter in India: recent trends and future outlook. *Air Qual. Atmos. Health* 12, 45–58. doi:10.1007/s11869-018-0629-6

Paredes-Belmar, G., Bronfman, A., Marianov, V., and Latorre-Núñez, G. (2017). Hazardous materials collection with multiple-product loading. *J. Clean. Prod.* 141, 909–919. doi:10.1016/j.jclepro.2016.09.163

Pasika, S., and Gandla, S. T. (2020). Smart water quality monitoring system with cost-effective using IoT. *Heliyon* 6 (7), e04096. doi:10.1016/j.heliyon.2020.e04096

Perez Santin, E., Rodríguez Solana, R., González García, M., García Suárez, M. D. M., Blanco Díaz, G. D., Cima Cabal, M. D., et al. (2021). Toxicity prediction based on artificial intelligence: a multidisciplinary overview. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 11 (5), e1516. doi:10.1002/wcms.1516

Perlman, A., Sacks, R., and Barak, R. (2014). Hazard recognition and risk perception in construction. *Saf. Sci.* 64, 22–31. doi:10.1016/j.ssci.2013.11.019

Pope, P., Van Eeckhout, E., Rofer, C., Baldridge, S., Ferguson, J., Jiracek, G., et al. (1996). Environmental waste site characterization utilizing aerial photographs. *remote Sens. Surf. Geophys.*

Pouyanfar, N., Harofte, S. Z., Soltani, M., Siavashy, S., Asadian, E., Ghorbani-Bidkorbeh, F., et al. (2022). Artificial intelligence-based microfluidic platforms for the sensitive detection of environmental pollutants: recent advances and prospects. *Trends Environ. Anal. Chem.* 34, e00160. doi:10.1016/j.teac.2022.e00160

Praveenchandar, J., Vetrithangam, D., Kaliappan, S., Karthick, M., Pegada, N. K., Patil, P. P., et al. (2022). IoT-based harmful toxic gases monitoring and fault detection on the sensor dataset using deep learning techniques. *Sci. Program.* 2022, 1–11. doi:10.1155/2022/7516328

Pule, M., Yahya, A., and Chuma, J. (2017). Wireless sensor networks: a survey on monitoring water quality. *J. Appl. Res. Technol.* 15 (6), 562–570. doi:10.1016/j.jart.2017.07.004

Rajan, R., Robin, D. T., and Vandanarani, M. (2019). Biomedical waste management in Ayurveda hospitals–current practices and future prospectives. *J. Ayurveda Integr. Med.* 10 (3), 214–221. doi:10.1016/j.jaim.2017.07.011

Rani, L., Thapa, K., Kanojia, N., Sharma, N., Singh, S., Grewal, A. S., et al. (2021). An extensive review on the consequences of chemical pesticides on human health and environment. *J. Clean. Prod.* 283, 124657. doi:10.1016/j.jclepro.2020.124657

Ravi, N., and Johnson, D. P. (2021). Artificial intelligence based monitoring system for onsite septic systems failure. *Process Saf. Environ. Prot.* 148, 1090–1097. doi:10.1016/j.psep.2021.01.049

Restas, A. (2015). Drone applications for supporting disaster management. *World J. Eng. Technol.* 3 (03), 316–321. doi:10.4236/wjet.2015.33c047

Ristic, B., Angley, D., Moran, B., and Palmer, J. L. (2017). Autonomous multi-robot search for a hazardous source in a turbulent environment. *Sensors* 17 (4), 918. doi:10.3390/s17040918

Ryu, J. H. (2022). UAS-based real-time water quality monitoring, sampling, and visualization platform (UASWQP). *HardwareX* 11, e00277. doi:10.1016/j.ohx.2022.e00277

Sahil, K., Mehta, P., Bhardwaj, S. K., and Dhaliwal, L. K. (2023). "Development of mitigation strategies for the climate change using artificial intelligence to attain sustainability," in *Visualization techniques for climate change with machine learning and artificial intelligence* (Elsevier), 421–448.

Sanaeifar, A., Zhang, W., Chen, H., Zhang, D., Li, X., and He, Y. (2022). Study on effects of airborne Pb pollution on quality indicators and accumulation in tea plants using Vis-NIR spectroscopy coupled with radial basis function neural network. *Ecotoxicol. Environ. Saf.* 229, 113056. doi:10.1016/j.ecoenv.2021.113056

Sari, M., Cosgun, T., Yalcin, I. E., Taner, M., and Ozyigit, I. I. (2022). Deciding heavy metal levels in soil based on various ecological information through artificial intelligence modeling. *Appl. Artif. Intell.* 36 (1), 2014189. doi:10.1080/08839514.2021.2014189

Sarkar, A., and Pandey, P. (2015). River water quality modelling using artificial neural network technique. *Aquat. Procedia* 4, 1070–1077. doi:10.1016/j.aqpro.2015.02.135

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., and Biancone, P. (2021). The role of artificial intelligence in healthcare: a structured literature review. *BMC Med. Inf. Decis. Mak.* 21, 125–223. doi:10.1186/s12911-021-01488-9

Seiber, C., Nowlin, D., Landowski, B., and Tolentino, M. E. (2018). "Tracking hazardous aerial plumes using IoT-enabled drone swarms," in 2018 IEEE 4th World Forum on Internet of Things (WF-IoT).

Seker, S. (2022). IoT based sustainable smart waste management system evaluation using MCDM model under interval-valued q-rung orthopair fuzzy environment. *Technol. Soc.* 71, 102100. doi:10.1016/j.techsoc.2022.102100

Seo, M., and Lee, S. W. (2022). Methodology to classify hazardous compounds via deep learning based on convolutional neural networks. *Curr. Appl. Phys.* 41, 59–65. doi:10.1016/j.cap.2022.06.003

Shafique, A., Cao, G., Khan, Z., Asad, M., and Aslam, M. (2022). Deep learning-based change detection in remote sensing images: a review. *Remote Sens.* 14 (4), 871. doi:10.3390/rs14040871

Sharma, P., Bano, A., Singh, S. P., Sharma, S., Xia, C. L., Nadda, A. K., et al. (2022). Engineered microbes as effective tools for the remediation of polyaromatic aromatic hydrocarbons and heavy metals. *Chemosphere* 306, 135538. doi:10.1016/j.chemosphere.2022.135538

Sharma, V., Arya, V. M., and Sharma, K. (2017). Spatial variability of soil pH and electrical conductivity in the Jammu district of Jammu and Kashmir. *J. Soil Water Conservation* 16 (4), 320–324. doi:10.5958/2455-7145.2017.00047.9

Shen, Y. (2018). Data sharing practices, information exchange behaviors, and knowledge discovery dynamics: a study of natural resources and environmental scientists. *Environ. Syst. Res.* 6 (1), 9–14. doi:10.1186/s40068-017-0086-5

Shi, L., Wang, X., Zhang, T., Hu, C., Luo, K., and Bai, B. (2016). Hazardous gas detection four-rotor UAV system development. 2016 IEEE International Conference on Mechatronics and Automation.

Shi, S., Hou, M., Gu, Z., Jiang, C., Zhang, W., Hou, M., et al. (2022). Estimation of heavy metal content in soil based on machine learning models. *Land* 11 (7), 1037. doi:10.3390/land11071037

Shirkoohi, M. G., Tyagi, R. D., Vanrolleghem, P. A., and Drogui, P. (2022). Artificial intelligence techniques in electrochemical processes for water and wastewater treatment: a review. *J. Environ. Health Sci. Eng.* 20 (2), 1089–1109. doi:10.1007/s40201-022-00835-w

Signes-Pastor, A. J., Carey, M., Carbonell-Barrachina, A. A., Moreno-Jiménez, E., Green, A. J., and Meharg, A. A. (2016). Geographical variation in inorganic arsenic in paddy field samples and commercial rice from the Iberian Peninsula. *Food Chem.* 202, 356–363. doi:10.1016/j.foodchem.2016.01.117

Singh, P., and Kaur, A. (2022). A systematic review of artificial intelligence in agriculture. *Deep Learn. Sustain. Agric.*, 57–80. doi:10.1016/b978-0-323-85214-2.00011-2

Slonecker, E., Lacerte, M., and Garofalo, D. (2002). The value of historic imagery. *Earth Obs. Mag.* 8, 39–41.

Slonecker, T., Fisher, G. B., Aiello, D. P., and Haack, B. (2010). Visible and infrared remote imaging of hazardous waste: a review. *Remote Sens.* 2 (11), 2474–2508. doi:10.3390/rs2112474

Sonne, C., Jenssen, B. M., Rinklebe, J., Lam, S. S., Hansen, M., Bossi, R., et al. (2023). EU need to protect its environment from toxic per-and polyfluoroalkyl substances. *Sci. Total Environ.* 876, 162770. doi:10.1016/j.scitotenv.2023.162770

Talebkeikhah, F., Rasam, S., Talebkeikhah, M., Torkashvand, M., Salimi, A., and Moraveji, M. K. (2022). Investigation of effective processes parameters on lead (II) adsorption from wastewater by biochar in mild air oxidation pyrolysis process. *Int. J. Environ. Anal. Chem.* 102 (16), 3975–3995. doi:10.1080/03067319.2020.1777291

Tan, K., Ma, W., Chen, L., Wang, H., Du, Q., Du, P., et al. (2021). Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning. *J. Hazard. Mater.* 401, 123288. doi:10.1016/j.jhazmat.2020.123288

Tawabini, B., Yassin, M. A., Benaafi, M., Adetoro, J. A., Al-Shaibani, A., and Abba, S. (2022). Spatiotemporal variability assessment of trace metals based on subsurface water quality impact integrated with artificial intelligence-based modeling. *Sustainability* 14 (4), 2192. doi:10.3390/su14042192

Thomas, G. W., Sousan, S., Tatum, M., Liu, X., Zuidema, C., Fitzpatrick, M., et al. (2018). Low-cost, distributed environmental monitors for factory worker health. *Sensors* 18 (5), 1411. doi:10.3390/s18051411

Trasande, L., Massey, R. I., DiGangi, J., Geiser, K., Olanipekun, A. I., and Gallagher, L. (2011). How developing nations can protect children from hazardous chemical exposures while sustaining economic growth. *Health Aff.* 30 (12), 2400–2409. doi:10.1377/hlthaff.2010.1217

Tsitsimpelis, I., Taylor, C. J., Lennox, B., and Joyce, M. J. (2019). A review of ground-based robotic systems for the characterization of nuclear environments. *Prog. Nucl. energy* 111, 109–124. doi:10.1016/j.pnucene.2018.10.023

Vasudevan, S. K., and Baskaran, B. (2021). An improved real-time water quality monitoring embedded system with IoT on unmanned surface vehicle. *Ecol. Inf.* 65, 101421. doi:10.1016/j.ecoinf.2021.101421

Vijayakumar, N., and Ramya, (2015). "The real time monitoring of water quality in IoT environment," in 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).

Vincent, T. A., Xing, Y., Cole, M., and Gardner, J. W. (2019). Investigation of the response of high-bandwidth MOX sensors to gas plumes for application on a mobile robot in hazardous environments. *Sensors Actuators B Chem.* 279, 351–360. doi:10.1016/j.snb.2018.08.125

Visvanathan, R., Kamarudin, K., Mamduh, S., Yeon, A., Zakaria, A., Kamarudin, L., et al. (2018). Gas sensing mobile robot: a review. *J. Telecommun. Electron. Comput. Eng. (JTEC)* 10 (1-15), 101–105.

Von Meding, J., Oyedele, L., and Cleland, D. (2009). Developing NGO competencies in post-disaster reconstruction: a theoretical framework. *Disaster Adv.* 2 (3), 36–45.

Wandel, M., Lilienthal, A. J., Duckett, T., Weimar, U., and Zell, A. (2003). "Gas distribution in unventilated indoor environments inspected by a mobile robot," in IEEE international conference on advanced robotics, ICAR 2003, Coimbra, Portugal, June 30-July 3, 2003.

Wang, B., Yun, M., Liu, Q., and Wang, Y. (2023a). Precision safety management (PSM): a novel and promising approach to safety management in the precision era. *Saf. Sci.* 157, 105931. doi:10.1016/j.ssci.2022.105931

Wang, C. J., Ng, C. Y., and Brook, R. H. (2020). Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. *Jama* 323 (14), 1341–1342. doi:10.1001/jama.2020.3151

Wang, J., Huang, Y., Zhai, W., Li, J., Ouyang, S., Gao, H., et al. (2023b). Research on coal mine safety management based on digital twin. *Heliyon* 9 (3), e13608. doi:10.1016/j.heliyon.2023.e13608

Wani, O. A., Akhter, F., Kumar, S. S., Babu, S., Kanth, R. H., Mir, S. A., et al. (2023). Mitigating soil erosion through biomass-derived biochar: exploring the influence of feedstock types and pyrolysis temperature. *Land* 12 (12), 2111. doi:10.3390/land12122111

Whitelaw, S., Mamas, M. A., Topol, E., and Van Spall, H. G. (2020). Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digital Health* 2 (8), e435–e440. doi:10.1016/s2589-7500(20)30142-4

Wilson, A. D. (2012). Review of electronic-nose technologies and algorithms to detect hazardous chemicals in the environment. *Procedia Technol.* 1, 453–463. doi:10.1016/j.protcy.2012.02.101

Wong, C., Yang, E., Yan, X.-T., and Gu, D. (2018). Autonomous robots for harsh environments: a holistic overview of current solutions and ongoing challenges. *Syst. Sci. Control Eng.* 6 (1), 213–219. doi:10.1080/21642583.2018.1477634

Wu, Y., Zhang, X., Liao, Q., and Ji, J. (2011). Can contaminant elements in soils be assessed by remote sensing technology: a case study with simulated data. *Soil Sci.* 176 (4), 196–205. doi:10.1097/ss.0b013e3182114717

Yan, H., Wan, J., Zhang, C., Tang, S., Hua, Q., and Wang, Z. (2018). Industrial big data analytics for prediction of remaining useful life based on deep learning. *IEEE access* 6, 17190–17197. doi:10.1109/access.2018.2809681

Yang, C.-T., Chen, H.-W., Chang, E.-J., Kristiani, E., Nguyen, K. L. P., and Chang, J.-S. (2021). Current advances and future challenges of AIoT applications in particulate matters (PM) monitoring and control. *J. Hazard. Mater.* 419, 126442. doi:10.1016/j.jhazmat.2021.126442

Yang, H., Yang, X., Ning, Z., Kwon, S. Y., Li, M.-L., Tack, F. M., et al. (2022). The beneficial and hazardous effects of selenium on the health of the soil-plant-human system: an overview. *J. Hazard. Mater.* 422, 126876. doi:10.1016/j.jhazmat.2021.126876

Yetilmezsoy, K., and Demirel, S. (2008). Artificial neural network (ANN) approach for modeling of Pb (II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells. *J. Hazard. Mater.* 153 (3), 1288–1300. doi:10.1016/j.jhazmat.2007.09.092

Young, S., Balluz, L., and Malilay, J. (2004). Natural and technologic hazardous material releases during and after natural disasters: a review. *Sci. Total Environ.* 322 (1-3), 3–20. doi:10.1016/s0048-9697(03)00446-7

Yusuf, A., Amusa, H. K., Eniola, J. O., Giwa, A., Pikuda, O., Dindi, A., et al. (2023). Hazardous and emerging contaminants removal from water by plasma-based treatment: a review of recent advances. *Chem. Eng. J. Adv.* 14, 100443. doi:10.1016/j.ceja.2023.100443

Zafar, M., Van Vinh, N., Behera, S. K., and Park, H.-S. (2017). Ethanol mediated as (III) adsorption onto Zn-loaded pinecone biochar: experimental investigation, modeling, and optimization using hybrid artificial neural network-genetic algorithm approach. *J. Environ. Sci.* 54, 114–125. doi:10.1016/j.jes.2016.06.008

Zahed, Z., Mufti, S., Kumar, S. S., Wani, O. A., Mushtaq, F., Rasool, R., et al. (2022). Organic and inorganic mulches combination improves the productivity, quality and profitability of rainfed potato in the temperate himalayan region. *Gesunde Pflanz.* 74 (4), 1109–1122. doi:10.1007/s10343-022-00650-1

Zeba, G., Dabić, M., Čičak, M., Daim, T., and Yalcin, H. (2021). Technology mining: artificial intelligence in manufacturing. *Technol. Forecast. Soc. Change* 171, 120971. doi:10.1016/j.techfore.2021.120971

Zhai, A., Hou, B., Ding, X., and Huang, G. (2021). Hazardous chemical accident prediction for drinking water sources in Three Gorges Reservoir. *J. Clean. Prod.* 296, 126529. doi:10.1016/j.jclepro.2021.126529

Zhang, Y. (2019). Research on key technologies of remote design of mechanical products based on artificial intelligence. *J. Vis. Commun. Image Represent.* 60, 250–257. doi:10.1016/j.jvcir.2019.02.010

Zhao, D., Wang, J., Jiang, X., Zhen, J., Miao, J., Wang, J., et al. (2022). Reflectance spectroscopy for assessing heavy metal pollution indices in mangrove sediments using XGBoost method and physicochemical properties. *Catena* 211, 105967. doi:10.1016/j.catena.2021.105967

Zhao, P., Gao, Y., and Sun, X. (2023). The impact of artificial intelligence on pollution emission intensity—evidence from China. *Environ. Sci. Pollut. Res.* 30 (39), 91173–91188. doi:10.1007/s11356-023-28866-2

Zhou, H., Yu, X., Alhaskawi, A., Dong, Y., Wang, Z., Jin, Q., et al. (2022). A deep learning approach for medical waste classification. *Sci. Rep.* 12 (1), 2159. doi:10.1038/s41598-022-06146-2

Zhu, X., Wang, X., and Ok, Y. S. (2019). The application of machine learning methods for prediction of metal sorption onto biochars. *J. Hazard. Mater.* 378, 120727. doi:10.1016/j.jhazmat.2019.06.004

Zinnuraain, S., Hasan, M., Hakque, M. A., and Arefin, M. M. N. (2019). "Smart gas leakage detection with monitoring and automatic safety system," in 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET).

Glossary

AAA	artificial algae algorithm
AEC	Architecture, Engineering, and Construction
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
CARS	Competitive Adaptive Reweighted Sampling
CNN	Convolutional Neural Network
ENN	Elman Neural Network
ERF	Extreme Random Forest
GHG	Green House Gas
GIS	Geographic Information System
GOA	Grasshopper Optimization Algorithm
GPS/GSM	Global Positioning System/Global System for Mobile Communication
HIS	Hyperspectral imaging
HRAI	High Resolution Aerial Imagery
IoT	Internet of Things
LED	Light-Emitting Diode
MCU	Wireless Microcontroller Unit
ML	Machine Learning
MLP	Multilayer Perception
MOX	Metal Oxide Semiconductor
MRO	Mobile Robotic Olfaction
MSS	Multispectral Sensors
MW	Medical Waste
NAV	Nano-Air Vehicle
NIR	Near-Infrared
PM	Particulate Matter
RF	Random Forest
RFID	Radio Frequency Identification
RNN	Recurrent Neural Network
SAR	Synthetic Aperture Radar
SERS	surface-enhanced Raman scattering
SMS	Short Message Service
SVM	Support Vector Machine
UAS	Unmanned Aerial Systems
UASWQP	Unmanned Aerial Systems Water Quality Platform
UAV	Unmanned Aerial Vehicles
WSN	Wireless Sensor Network

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 02 April 2024
DOI 10.3389/fenvs.2024.1243962

OPEN ACCESS

EDITED BY
Sayali Sandbhor,
Symbiosis International University, India

REVIEWED BY
Komal Shukla,
Environment and Climate Change Canada (ECCC), Canada
Roseline Oluwaseun Ogundokun,
Landmark University, Nigeria

*CORRESPONDENCE
Bian Chao,
15526387@qq.com

RECEIVED 21 June 2023
ACCEPTED 14 February 2024
PUBLISHED 02 April 2024

CITATION
Chao B and Guang Qiu H (2024), Air pollution concentration fuzzy evaluation based on evidence theory and the K-nearest neighbor algorithm.
Front. Environ. Sci. 12:1243962.
doi: 10.3389/fenvs.2024.1243962

Air pollution concentration fuzzy evaluation based on evidence theory and the K-nearest neighbor algorithm

Bian Chao[1,2]* and Huang Guang Qiu[1]

[1]Management College Xi'an University of Architecture and Technology, Xi'an, China, [2]Yinchuan Institute of Science and Technology, Yinchuan, China

Background: Air pollution, characterized by complex spatiotemporal dynamics and inherent uncertainty, poses significant challenges in accurate air quality prediction, and current methodologies often fail to adequately address these complexities.

Objective: This study presents a novel fuzzy modeling approach for estimating air pollution concentrations.

Methods: This fuzzy evaluation method integrates an improved evidence theory with comprehensive weighting and the K-nearest neighbor (KNN) interval distance within the framework of the matter-element extension model. This involves generating the basic probability assignment (BPA) based on interval similarity, performing sequential fusion using the Dempster–Shafer evidence theory, enhancing the fusion results via comprehensive weighting, and conducting fuzzy evaluation of air pollution concentrations using the matter-element extension KNN interval distance.

Results: Our method achieved significant improvements in monitoring air pollution concentrations, incorporating spatiotemporal factors and pollutant concentrations more effectively than existing methods. Implementing sequential fusion and subjective–objective weighting reduced the error rate by 38% relative to alternative methods.

Discussion: Fusion of multi-source air pollution data via this method effectively mitigates inherent uncertainty and enhances the accuracy of the KNN method. It produces more comprehensive air pollution concentration fusion results, improving accuracy by considering spatiotemporal correlation, toxicity, and pollution levels. Compared to traditional air-quality indices, our approach achieves greater accuracy and better interpretability, making it possible to develop more effective air quality management strategies. Future research should focus on expanding the dataset to include more diverse geographical and meteorological conditions, further refining the model to integrate external factors like meteorological data and regional industrial activity, and improving computational efficiency for real-time applications.

KEYWORDS

evidence theory, K-nearest neighbor, air pollution concentration fusion, fuzzy evaluation method, basic probability assignment

Abbreviations: AQI, air quality index; BPA, basic probability assignment.

1 Introduction

Atmospheric pollution is an urgent issue globally, especially in regions undergoing rapid industrialization and urbanization. The detrimental effects of pollution on ecological and climatic stability and human health are increasingly evident (Murena, 2004; Chen and Zhu, 2014; Chen et al., 2022). With the increasing global focus on environmental protection and low-carbon development, research is focusing on how technological innovation, green finance, and policy design can reduce environmental pollution and foster sustainable development. For example, Li et al. (2023) explored how digital finance could facilitate green technological innovations in polluting industries by easing funding constraints and augmenting research and development investments. Meanwhile, Feng et al. (2023) focused on reducing NO_x emissions via advanced catalytic technologies, thus improving air quality. The intensification of global climate change has amplified the importance of research into urban heat-island effects (Shang et al., 2023), biodiversity loss (Wang et al., 2022), and the global carbon cycle (Zhang et al., 2021; Xiong et al., 2022).

Within this context, Wu et al. (2023) investigated the impact of clustered institutional investors on low-carbon innovation in family businesses, discussing how green finance and family governance can promote sustainable development in a changing economic environment. Kong et al. (2023) proposed a lifecycle-oriented low-carbon product design method to meet the challenges of climate change. Using cloud computing technology, Shang et al. (2021) explored factors influencing urban carbon footprints in China, and proposed key strategies for optimizing carbon emission predictions and low-carbon economic development (Luo et al., 2024). These studies provide technological and financial solutions while highlighting the significance of policy design in advancing environmental protection and sustainable development, which require accurate assessment of atmospheric pollution.

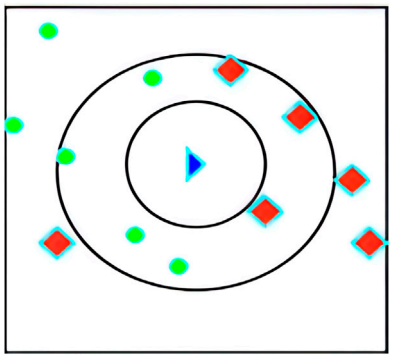


FIGURE 1
The K-nearest neighbor (KNN) model. Blue: sample to be classified; red: classified sample; green: classified sample.

The air quality index (AQI) is a widely accepted assessment tool for air pollution; however, it presents several limitations. For instance, by aggregating the concentration levels of multiple

TABLE 1 Literature review of related studies.

References	Method	Advantage	Disadvantage
Murena (2004)	Involved developing an air quality index (AQI) through principal component analysis and linear algebra to rank states by pollution levels based on data from five pollutants in the transportation sector	Furnished an initial understanding of the impact of air pollution in industrialized and urbanized areas	Possibly lacked granular data and in-depth analysis of complex pollutant interactions
Chen and Zhu (2014)	Delved deeper into the air pollution issues of industrialized and urbanized areas	Focused on high-risk areas, delivering more specific data on atmospheric pollution	Centered on particular regions, limiting the broader applicability of the research findings
Pope III and Dockery (2006)	Leveraged more precise data and measurement methodologies to study atmospheric pollution	Provided high-quality data, advancing the accuracy of air pollution research	Used many resources and technology to acquire and analyze data
Vaidya and Kumar (2006)	Evaluated and critiqued the limitations and deficiencies of the analytic hierarchy process (AHP)	Exposed the constraints of AHP, promoting the development of new methodologies and techniques	Primarily represented a critical study, offering no explicit solutions
Li et al. (2010)	Investigated the performance of the grey relational analysis when dealing with intricate, non-linear environmental variables	Disclosed the method's applicability in certain contexts	Faced challenges in data integrity and precision when handling complex, non-linear environmental elements
Carslaw and Rhys-Tyler (2013)	Employed deep learning technologies to predict concentrations of NO_2 and PM_{10}	Enhanced prediction accuracy and adaptability, capable of processing intricate datasets	Required substantial data and computational resources
Cui et al. (2022)	Criticized the AQI for its shortcomings in handling complex atmospheric pollutants	Pointed out the limitations of AQI, offering a direction for the improvement and development of new assessment standards	Presented a critical study that failed to provide specific solutions or measures for improvement
Sun et al. (2022b)	Adopted the improved D-S evidence theory for a comprehensive assessment of air quality; determined the weights of evidence through the entropy weight method and introduced decision credibility by calculating the dispersion of different evidence decisions	Validated the efficacy of the DCre-Weight model, which showed higher accuracy and consistency than other enhanced methods of the evidence theory and recent fuzzy synthetic evaluation methods; yielded improved credibility of fusion results and well-articulated uncertainty	Presented limitations in complexity, applicability, and comparison with other latest methods

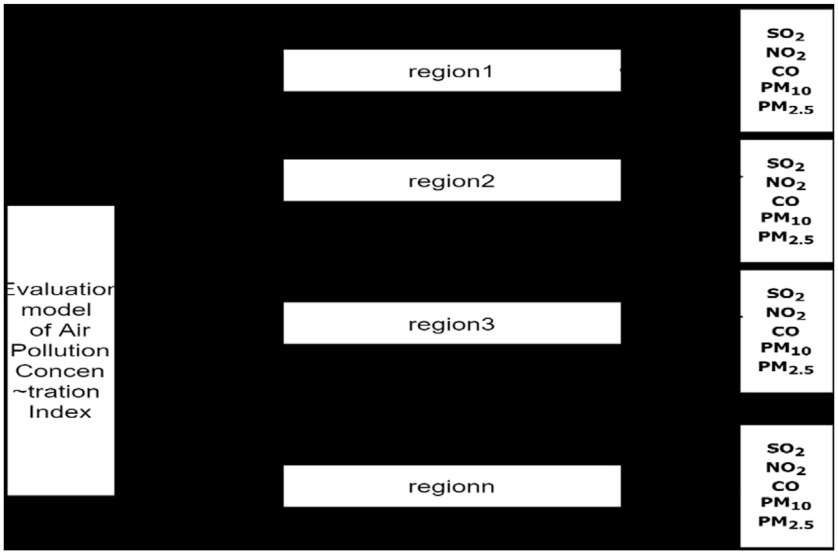


FIGURE 2
Pollution concentration evaluation factor set.

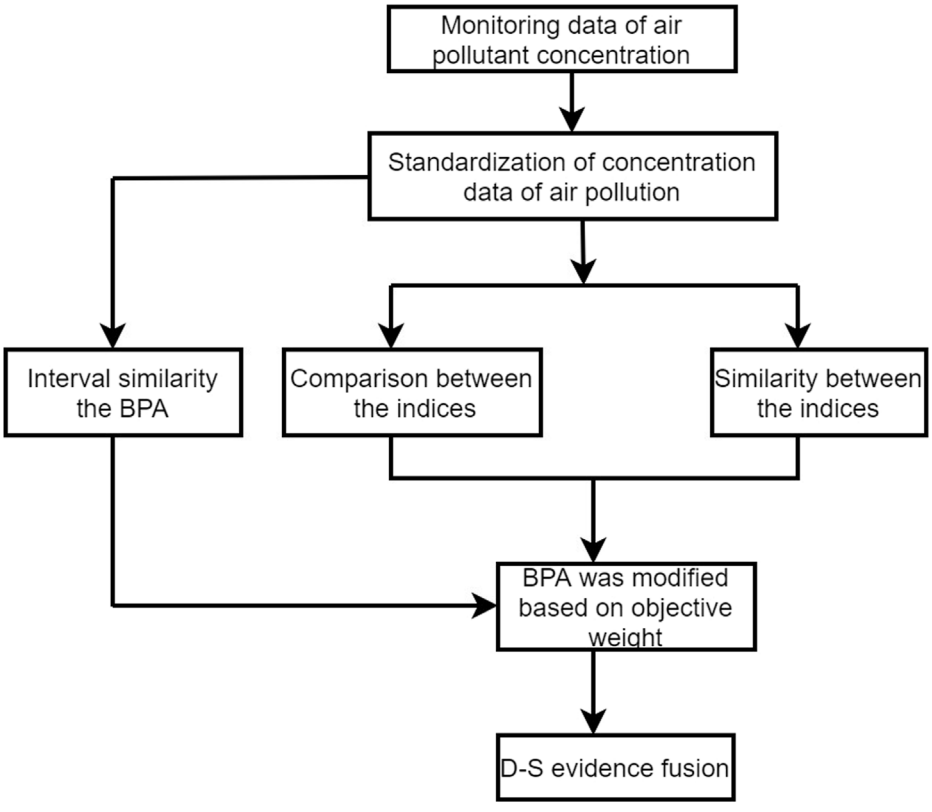


FIGURE 3
Air pollution concentration evaluation model.

pollutants into a single metric, the AQI may obscure the risks of high concentrations of specific pollutants (Pope et al., 2022). Further complications arise from the fact that not all pollutants are considered in AQI computations (Priti and Kumar, 2022), and the AQI calculations and standards differ between countries, leading to interpretative challenges (Karavas et al.,

TABLE 2 Air pollution concentration standard.

Pollutant	Air pollution concentration levels (μg/m^3)						
	I	II	III	IV	V	VI	VII
$PM_{2.5}$	35	75	115	150	250	350	500
PM_{10}	50	150	250	350	420	500	600
SO_2	50	150	475	800	1,600	2,100	2,620
NO_2	40	80	180	280	565	750	940
CO	2	4	14	24	36	48	60
$O3_8$	100	160	215	265	800	1,000	1,200

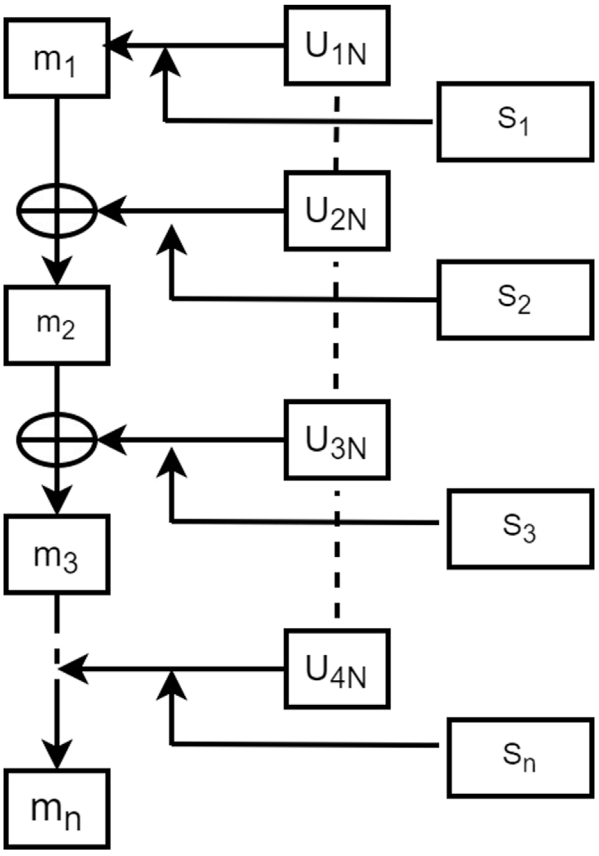


FIGURE 4
Sequential fusion via the Dempster–Shafer (DS) theory based on interval similarity.

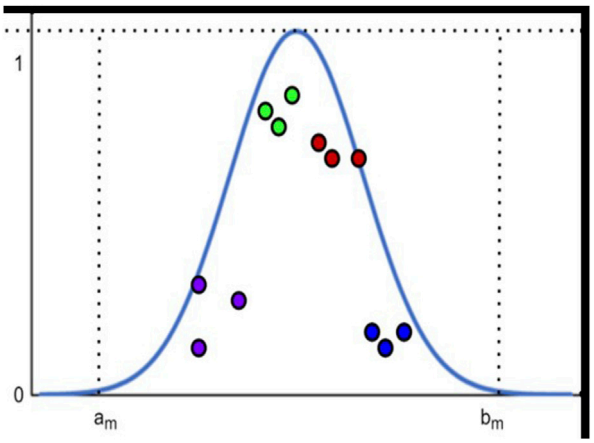


FIGURE 5
Matter-element extensible distance.

TABLE 3 Nondimensionalization of the air pollution concentration standards.

Pollutant	Air pollution concentration levels						
	I	II	III	IV	V	VI	VII
$PM_{2.5}$	0.17	0.36	0.55	0.71	1.19	1.66	2.37
PM_{10}	0.15	0.45	0.75	1.06	1.27	1.51	1.81
SO_2	0.04	0.13	0.43	0.72	1.44	1.89	2.35
NO_2	0.10	0.20	0.44	0.69	1.40	1.85	2.32
CO	0.07	0.15	0.52	0.89	1.34	1.79	2.23
$O3_8$	0.28	0.34	0.52	0.69	1.38	1.72	2.07

2021). Moreover, the AQI is inadequate for capturing and processing the complexities of atmospheric pollutants and their interactions (Cui et al., 2022). For instance, Zhang and Cao (2015) examined the heterogeneity and complexity of $PM_{2.5}$ levels, highlighting the shortcomings of current assessment strategies. Further, accurate data and measurement techniques are indispensable for controlling air pollution (Pope III and Dockery, 2006).

Vaidya and Kumar (2006) expressed concerns about the accuracy and consistency of other traditional assessment methods, such as the analytic hierarchy process, under complex, dynamic, and variable environmental conditions. While the analytic hierarchy process, a multi-criteria decision-making tool, is used to determine the significance of various pollutants (Saaty, 2008), its reliance on subjective expert judgments may introduce bias (Dyer and Forman, 1992). Grey relational analysis may be suited to specific scenarios, but it presents challenges in model selection and membership function when dealing with complex, non-linear environmental factors (Wang and Klir, 2009; Li et al., 2010). Quantification of uncertain information was proposed by Zadeh (1965), and the fuzzy comprehensive evaluation method (Mo et al., 2020) offered a new approach to handle ambiguous and uncertain data. However, challenges persist regarding the scientific rigor and reproducibility of data generated using the analytic hierarchy process.

Considering these issues, attention has shifted towards more advanced and innovative assessment techniques. The potential of evidence theory and K-nearest neighbor (KNN) algorithms to handle complex and dynamic air quality data has been widely

TABLE 4 Concentrations of different air pollutants in different regions.xi'an.

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
Xingqing District	11	70	1.3	21	99	64
Xiaozhai	12	76	1	26	96	63
Municipal People's Stadium	8	80	1.1	22	121	76
Kwong Wan Tam	11	66	1.2	20	118	70
Bureau of Culture and Sports	10	51	1.2	48	118	97
Radio monitoring center	20	59	1	27	132	98
Qujiang	8	62	1	24	114	60
Textile City	13	45	1.3	18	123	74
Economic Development Zone	12	74	1	26	110	83
Grass Beach (control point)	10	55	0.9	27	168	99
Chang'an District	9	52	0.9	44	95	62
High-voltage switch factory	11	66	0.9	27	120	84
High-tech Western District	11	71	1	25	99	77
Lintong District	15	67	1.2	34	97	67
Xingqing District	18	84	2	16	186	126
Xiaozhai	20	94	1.5	11	174	120
Municipal People's Stadium	18	91	1.7	21	199	128
Kwong Wan Tam	21	71	2	25	191	142
Bureau of Culture and Sports	16	70	1.5	28	192	152
Radio monitoring center	28	57	1.4	46	146	117
Qujiang	12	81	1.5	8	194	112
Textile City	19	56	2	12	194	138
Economic Development Zone	26	82	1.7	30	169	133
Grass Beach (control point)	19	66	1.6	27	203	129
Chang'an District	14	75	1.5	13	192	124
High-voltage switch factory	23	82	1.5	21	201	135
High-tech Western District	22	88	1.6	14	173	127
Lintong District	24	72	2.2	48	167	120
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Xingqing District	10	75	1.1	33	98	57
Xiaozhai	13	80	1.4	35	150	60
Municipal People's Stadium	11	82	1.4	41	122	60
Kwong Wan Tam	10	76	0.8	19	109	65
Bureau of Culture and Sports	11	55	0.9	63	106	61
Radio monitoring center	13	58	1.1	27	196	118

(Continued on following page)

TABLE 4 (*Continued*) Concentrations of different air pollutants in different regions.xi'an.

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
Qujiang	10	71	1.1	23	192	75
Textile City	8	71	1	22	113	62
Economic Development Zone	13	84	1.1	39	167	74
Grass Beach (control point)	8	68	1	29	149	85
Chang'an District	11	55	0.8	53	129	54
High-voltage switch factory	14	80	1.2	54	156	80
High-tech Western District	13	77	1.3	55	124	73
Lintong District	14	73	1	31	127	76

TABLE 5 Concentrations of different air pollutants in different regions.beijing.

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
Yanqing Summer Capital	9	31	0.8	49	69	31
Miyun New City	8	21	0.5	50	43	19
Pinggu New City	4	30	0.8	29	65	34
Miyun Town	7	38	0.7	36	54	23
Fengtai Xiaotun	8	64	1	21	104	39
Huairou New City	6	23	0.5	40	42	19
Yanqing Shiheying	9	32	0.9	46	62	32
Daxing Old Palace	4	61	0.7	16	80	33
Fangshan and Yanshan	6	36	0.8	48	56	31
Tongzhou Dongguan	5	57	0.6	15	85	36
Fengtai Yungang	7	46	0.6	22	66	29
Three stores in Mentougou	8	43	0.9	32	70	29
ancient city	6	47	0.8	20	95	38
Olympic Sports Center	7	54	0.6	25	61	27
Changping Town	8	34	0.7	26	52	25
Huairou Town	5	30	0.6	38	39	20
Shunyi New City	7	42	0.8	28	60	29
Haidian Wanliu	10	53	0.9	22	76	30
Official Garden	7	59	0.8	21	62	29
Agricultural Exhibition Hall	5	63	0.6	18	74	30
Temple of Heaven	3	58	0.9	15	58	32
Dongsi	4	52	0.6	18	63	30
Dingling (control point)	7	30	0.5	28	46	28
Wanshou West Palace	5	63	0.9	17	67	32

(Continued on following page)

TABLE 5 (*Continued*) Concentrations of different air pollutants in different regions.beijing.

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Yanqing Summer Capital	5	39	0.6	36	47	20
Miyun New City	2	68	0.8	25	68	26
Pinggu New City	3	19	0.3	46	25	15
Miyun Town	3	36	0.6	38	39	19
Fengtai Xiaotun	4	59	0.8	23	51	24
Huairou New City	5	37	0.5	37	35	20
Yanqing Shiheying	5	65	0.7	26	68	28
Daxing Old Palace	3	52	0.5	30	48	21
Fangshan and Yanshan	3	45	1.1	27	65	24
Tongzhou Dongguan	2	26	0.4	42	43	17
Fengtai Yungang	3	29	0.6	42	41	19
Three stores in Mentougou	4	37	0.6	37	40	21
ancient city	2	57	0.7	24	59	23
Olympic Sports Center	5	67	0.7	23	51	22
Changping Town	2	37	0.4	33	37	16
Huairou Town	3	25	0.4	41	30	15
Shunyi New City	3	47	0.7	35	42	20
Haidian Wanliu	4	65	0.7	19	57	21
Official Garden	6	65	0.7	22	46	21
Agricultural Exhibition Hall	5	67	0.8	20	61	29
Temple of Heaven	3	56	0.7	28	36	21
Dongsi	3	57	0.8	24	50	23
Dingling (control point)	2	21	0.4	42	29	14
Wanshou West Palace	6	71	0.9	23	49	23

explored (Cover and Hart, 1967; Dempster, 1967; Shafer, 1976; Xiao et al., 2013; Dai et al., 2018; Wang et al., 2021; Sun et al., 2022a; Franklin et al., 2023). However, their efficacy under various geographical and meteorological conditions remains questionable. The research of Carslaw and Rhys-Tyler (2013), and the application of deep learning in predicting the concentrations of pollutants such as NO_2 and PM_{10} (Kukkonen et al., 2003), signal a shift towards more accurate and flexible assessment techniques. As a significant improvement, Sun et al. (2022b) presented an ambient air quality evaluation model based on an advanced evidence theory. Important research developments in this field are presented in Table 1. While advanced air quality assessment techniques provide significant insights, they are often unable to distinguish between sources of pollution and may not account for the latest scientific understanding of pollutant interactions. Moreover, the evolving nature of industrial emissions and urban development requires more adaptable and sophisticated analytical methods. Thus, recognizing these shortcomings, the development and deployment of more precise and efficient air quality assessment techniques have become paramount for both scientific research and policy formulation.

In this study, we present a fuzzy evaluation method that innovatively integrates evidence theory and the KNN algorithm, aiming to enhance atmospheric pollution concentration assessment precision and efficiency. This approach strives to overcome the limitations of conventional evaluation methods by synthesizing multifaceted, uncertain, and ambiguous environmental data, thus improving the precision and reliability. Via comparison with established evaluation techniques, our study reveals the proposed

TABLE 6 Concentrations of different air pollutants in different regions.tianjin

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
BinshuiEastRoad	22	57	1.6	20	97	50
JiansheRoad	18	67	1.5	23	79	43
NorthRingRoad	10	36	1.2	45	56	29
Xisido	18	59	1.2	20	85	50
ZhongshanNorthRoad	14	64	1.4	14	95	49
DaliRoad	17	65	1.3	18	81	52
BinshuiWestRoad	15	57	1.2	14	84	52
JinguRoad	18	67	1.4	16	92	59
HexiYijingRoad	21	57	1.5	23	78	45
DiweiRoad	19	59	1.7	24	93	67
XinlaoRoad	20	74	1.9	19	102	66
YongyangWestRoad	15	61	1.5	20	90	47
Tuanpowa	10	64	1.4	18	93	59
HanbeiRoad	20	52	1.7	19	78	44
YongmingRoad	23	72	1.6	19	88	58
FourthStreet	20	64	1.9	14	77	53
LeapForwardRoad	14	68	1.5	20	99	57
HuaiheRoad	16	55	1	17	81	47
Forwardlane	19	66	1.4	20	94	56
DazhiguNo8Road	16	67	1.5	16	70	53
Thepathofdiligenceandfrugality	20	68	1.3	15	84	44
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Yanqing Summer Capital	18	46	0.8	41	54	22
Miyun New City	14	54	0.9	35	63	31
Pinggu New City	11	28	0.9	47	40	18
Miyun Town	12	60	0.8	32	52	25
Fengtai Xiaotun	14	70	1.1	35	82	34
Huairou New City	18	60	0.9	34	90	48
Yanqing Shiheying	12	63	0.9	37	73	33
Daxing Old Palace	29	67	0.8	31	72	34
Fangshan and Yanshan	15	54	1.1	38	59	28
Tongzhou Dongguan	9	59	1.1	38	57	27
Fengtai Yungang	14	57	1.1	40	57	25
Three stores in Mentougou	11	46	0.7	50	39	18
ancient city	12	65	0.6	37	71	31

(Continued on following page)

TABLE 6 (*Continued*) Concentrations of different air pollutants in different regions.tianjin

Region	Pollutant concentration					
	SO_2 (μg/m³)	NO_2 (μg/m³)	CO (mg/m³)	O_3 (μg/m³)	PM_{10} (μg/m³)	$PM_{2.5}$ (μg/m³)
Olympic Sports Center	9	47	0.5	48	38	16
Changping Town	14	58	0.8	40	62	30
Huairou Town	19	70	0.7	29	57	25
Shunyi New City	12	62	0.9	37	75	33
Haidian Wanliu	15	71	1.3	34	78	33
Official Garden	14	63	1.1	40	64	30
Agricultural Exhibition Hall	13	65	0.8	36	65	30
Temple of Heaven	16	67	1.3	40	70	31

method's advantages and contributions, which can be summarized as follows:

a) Holistic evaluation of atmospheric pollution concentrations: The approach introduces an assessment technique that overcomes the constraints of traditional AQI. It integrates multi-source data and physical property analyses, comprehensively evaluates atmospheric pollution patterns, and accurately quantifies key pollutant concentrations.
b) Innovative synthesis of evidence theory and KNN: This synthesis enhances the precision of data fusion and relevance of the results.
c) Focus on data quality and accuracy: This hybrid approach emphasizes data integrity by applying interval similarity and subjective–objective weighting and considering variability in pollutant concentrations and toxicological characteristics.

To validate the robustness of this approach and its applicability in diverse environmental contexts, the analysis includes data from representative days across three distinct urban settings: Xi'an, Beijing, and Tianjin, rather than focusing on a single urban area. These cities were chosen for their varied geographic and meteorological profiles, thus enhancing the robustness of the findings. The findings confirm the method's efficacy across varied urban environments and address complex regional differences in air pollution. This multifaceted approach substantially enhances methodological rigor in air quality assessment.

2 Materials and methods

2.1 Theoretical background

2.1.1 KNN method

The KNN method, a commonly used machine learning algorithm for classification and regression, stores all available cases and classifies new data or cases based on a similarity (distance) measure. It functions by finding the K-nearest neighbors to the unknown sample within the known samples (Altman, 1992; Huihui and Yanming, 2013; Wang et al., 2021). Next, the class of the unknown sample is determined based on the class of the nearest neighbors, typically via "majority rule." The three elements of the KNN model are the choice of k, a distance measurement, and a classification-decision rule.

Regarding the choice of k, the blue triangle in Figure 1 represents the sample to be classified. When $k = 7$, of the seven samples nearest the sample to be classified, three belong to the red class and four to the green class. Therefore, the sample to be classified is predicted to belong to the green class. However, when $k = 9$, the sample to be classified belongs to the red class. Therefore, when the samples are unbalanced, the choice of k substantially impacts the results.

Next, distance is measured as follows:

Let the feature space χ be an n-dimensional real vector space R^n, x_i and $x_j \in \chi$, $x_i = (x_{i1}, x_{i2}.....x_{in})^T$, $x_j = (x_{j1}, x_{j2}.....x_{jn})^T$, with the distance d_P between x_i and x_j defined as

$$d_P(x_i, x_j) = \left(\sum_{l=1}^{n} |x_{il} - x_{jl}|^2\right)^{\frac{1}{P}} \quad (1)$$

When $p = 2$, this is the Euclidean distance:

$$d_2(x_i, x_j) = \left(\sum_{l=1}^{n} |x_{il} - x_{jl}|^2\right)^{\frac{1}{2}} \quad (2)$$

when $p = 1$, it is the Manhattan distance:

$$d_1(x_i, x_j) = \left(\sum_{l=1}^{n} |x_{il} - x_{jl}|\right) \quad (3)$$

and when $p = \infty$, it is the maximum distance between the coordinates, i.e.,

$$d_\infty(x_i, x_j) = \max_l \left(\sum_{l=1}^{n} |x_{il} - x_{jl}|\right) \quad (4)$$

2.1.2 Evidence theory

The evidence theory, proposed by Dempster (2008), has been further developed by Shafer (Ai et al., 2022; He et al., 2022; Liu et al., 2022; Ren et al., 2022). Evidence theory applies fuzzy logic to handle uncertainty, via the following steps.

a) Establishment of a discernment framework

The discernment framework is a set of all objects or entities under consideration. Subsets in the discernment framework are defined as follows:

TABLE 7 Standardized air pollutant concentrations across different regions.xi'an.

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Xingqing District	0.25	0.71	1.00	0.10	0.05	0.10
Xiaozhai	0.33	0.89	0.25	0.27	0.01	0.08
Municipal People's Stadium	0.00	1.00	0.50	0.13	0.36	0.41
Kwong Wan Tam	0.25	0.60	0.75	0.07	0.32	0.26
Bureau of Culture and Sports	0.17	0.17	0.75	1.00	0.32	0.95
Radio monitoring center	1.00	0.40	0.25	0.30	0.51	0.97
Qujiang	0.00	0.49	0.25	0.20	0.26	0.00
Textile City	0.42	0.00	1.00	0.00	0.38	0.36
Economic Development Zone	0.33	0.83	0.25	0.27	0.21	0.59
Grass Beach (control point)	0.17	0.29	0.00	0.30	1.00	1.00
Chang'an District	0.08	0.20	0.00	0.87	0.00	0.05
High-voltage switch factory	0.25	0.60	0.00	0.30	0.34	0.62
High-tech Western District	0.25	0.74	0.25	0.23	0.05	0.44
Lintong District	0.58	0.63	0.75	0.53	0.03	0.18
Xingqing District	0.38	0.74	0.75	0.20	0.70	0.35
Xiaozhai	0.50	1.00	0.13	0.08	0.49	0.20
Municipal People's Stadium	0.38	0.92	0.38	0.33	0.93	0.40
Kwong Wan Tam	0.56	0.39	0.75	0.43	0.79	0.75
Bureau of Culture and Sports	0.25	0.37	0.13	0.50	0.81	1.00
Radio monitoring center	1.00	0.03	0.00	0.95	0.00	0.13
Qujiang	0.00	0.66	0.13	0.00	0.84	0.00
Textile City	0.44	0.00	0.75	0.10	0.84	0.65
Economic Development Zone	0.88	0.68	0.38	0.55	0.40	0.53
Grass Beach (control point)	0.44	0.26	0.25	0.48	1.00	0.43
Chang'an District	0.13	0.50	0.13	0.13	0.81	0.30
High-voltage switch factory	0.69	0.68	0.13	0.33	0.96	0.58
High-tech Western District	0.63	0.84	0.25	0.15	0.47	0.38
Lintong District	0.75	0.42	1.00	1.00	0.37	0.20
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Xingqing District	0.33	0.69	0.50	0.32	0.00	0.05
Xiaozhai	0.83	0.86	1.00	0.36	0.53	0.09
Municipal People's Stadium	0.50	0.93	1.00	0.50	0.24	0.09
Kwong Wan Tam	0.33	0.72	0.00	0.00	0.11	0.17
Bureau of Culture and Sports	0.50	0.00	0.17	1.00	0.08	0.11
Radio monitoring center	0.83	0.10	0.50	0.18	1.00	1.00

(Continued on following page)

TABLE 7 (*Continued*) Standardized air pollutant concentrations across different regions.xi'an.

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Qujiang	0.33	0.55	0.50	0.09	0.96	0.33
Textile City	0.00	0.55	0.33	0.07	0.15	0.13
Economic Development Zone	0.83	1.00	0.50	0.45	0.70	0.31
Grass Beach (control point)	0.00	0.45	0.33	0.23	149	85
Chang'an District	0.50	0.00	0.00	0.77	129	54
High-voltage switch factory	1.00	0.86	0.67	0.80	156	80
High-tech Western District	0.83	0.76	0.83	0.82	124	73
Lintong District	1.00	0.62	0.33	0.27	127	76

TABLE 8 Standardized air pollutant concentrations across different regions.beijing

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
Yanqing Summer Capital	0.86	0.23	0.60	0.97	0.46	0.60
Miyun New City	0.71	0.00	0.00	1.00	0.06	0.00
Pinggu New City	0.14	0.21	0.60	0.40	0.40	0.75
Miyun Town	0.57	0.40	0.40	0.60	0.23	0.20
Fengtai Xiaotun	0.71	1.00	1.00	0.17	1.00	1.00
Huairou New City	0.43	0.05	0.00	0.71	0.05	0.00
Yanqing Shiheying	0.86	0.26	0.80	0.89	0.35	0.65
Daxing Old Palace	0.14	0.93	0.40	0.03	0.63	0.70
Fangshan and Yanshan	0.43	0.35	0.60	0.94	0.26	0.60
Tongzhou Dongguan	0.29	0.84	0.20	0.00	0.71	0.85
Fengtai Yungang	0.57	0.58	0.20	0.20	0.42	0.50
Three stores in Mentougou	0.71	0.51	0.80	0.49	0.48	0.50
ancient city	0.43	0.60	0.60	0.14	0.86	0.95
Olympic Sports Center	0.57	0.77	0.20	0.29	0.34	0.40
Changping Town	0.71	0.30	0.40	0.31	0.20	0.30
Huairou Town	0.29	0.21	0.20	0.66	0.00	0.05
Shunyi New City	0.57	0.49	0.60	0.37	0.32	0.50
Haidian Wanliu	1.00	0.74	0.80	0.20	0.57	0.55
Official Garden	0.57	0.88	0.60	0.17	0.35	0.50
Agricultural Exhibition Hall	0.29	0.98	0.20	0.09	0.54	0.55
Temple of Heaven	0.00	0.86	0.80	0.00	0.29	0.65
Dongsi	0.14	0.72	0.20	0.09	0.37	0.55
Dingling (control point)	0.57	0.21	0.00	0.37	0.11	0.45
Wanshou West Palace	0.29	0.98	0.80	0.06	0.43	0.65
:	:	:	:	:	:	:

(Continued on following page)

TABLE 8 (*Continued*) Standardized air pollutant concentrations across different regions.beijing

Region	Pollutant concentration					
	SO_2 (μg/m³)	NO_2 (μg/m³)	CO (mg/m³)	O_3 (μg/m³)	PM_{10} (μg/m³)	$PM_{2.5}$ (μg/m³)
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Yanqing Summer Capital	0.75	0.38	0.38	0.63	0.51	0.40
Miyun New City	0.00	0.94	0.63	0.22	1.00	0.80
Pinggu New City	0.25	0.00	0.00	1.00	0.00	0.07
Miyun Town	0.25	0.33	0.38	0.70	0.33	0.33
Fengtai Xiaotun	0.50	0.77	0.63	0.15	0.60	0.67
Huairou New City	0.75	0.35	0.25	0.67	0.23	0.40
Yanqing Shiheying	0.75	0.88	0.50	0.26	1.00	0.93
Daxing Old Palace	0.25	0.63	0.25	0.41	0.53	0.47
Fangshan and Yanshan	0.25	0.50	1.00	0.30	0.93	0.67
Tongzhou Dongguan	0.00	0.13	0.13	0.85	0.42	0.20
Fengtai Yungang	0.25	0.19	0.38	0.85	0.37	0.33
Three stores in Mentougou	0.50	0.35	0.38	0.67	0.35	0.47
ancient city	0.00	0.73	0.50	0.19	0.79	0.60
Olympic Sports Center	0.75	0.92	0.50	0.15	0.60	0.53
Changping Town	0.00	0.35	0.13	0.52	0.28	0.13
Huairou Town	0.25	0.12	0.13	0.81	0.12	0.07
Shunyi New City	0.25	0.54	0.50	0.59	0.40	0.40
Haidian Wanliu	0.50	0.88	0.50	0.00	0.74	0.47
Official Garden	1.00	0.88	0.50	0.11	0.49	0.47
Agricultural Exhibition Hall	0.75	0.92	0.63	0.04	0.84	1.00
Temple of Heaven	0.25	0.71	0.50	0.33	0.26	0.47
Dongsi	0.25	0.73	0.63	0.19	0.58	0.60
Dingling (control point)	0.00	0.04	0.13	0.85	0.09	0.00
Wanshou West Palace	1.00	1.00	0.75	0.15	0.56	0.60

$$\Theta = \{\theta_1, \theta_2, \theta_3 \cdots\cdots \theta_N\} \tag{5}$$

If $\boldsymbol{\Theta}$ is a finite complete set of N mutually exclusive elements, it is called a discernment framework. The set of $2N$ elements formed by the power set 2^{Θ} of Θ is:

$$2^{\theta} = \{\varphi, \theta_1, \theta_2, \theta_3 \cdots\cdots \theta_N, \theta_1 \cup \theta_2 \cup \theta_3 \cdots, \Theta\} \tag{6}$$

b) Determination of the basic probability assignment (BPA)

Initially, the evidence-based confidence (support) in each combination is established by the evidence processor. The basic probability is assigned as follows:

Let Θ be a discernment framework. The power set 2^{Θ} of Θ forms the set of propositions 2^{Θ}, $\forall A \in \Theta$.

If the function m: $2^{\Theta} \rightarrow [0, 1]$ satisfies the conditions

$$m(\phi) = 0 \text{ and } \sum_{A \subseteq \Theta} m(A) = 1 \tag{7}$$

then m is referred to as the BPA, and m(A) is the basic probability assigned to proposition A (i.e., the confidence accurately assigned to A).

There are two ways to calculate the BPA: via expert opinion or by constructing the corresponding mathematical model. The approach of using expert experience requires values provided by different experts, and thus tends to be subjective. Therefore, this method calculates BPA by constructing the corresponding mathematical model.

c) Evidence-combination rule

TABLE 9 Standardized air pollutant concentrations across different regions.tianjing

Region	Pollutant concentration					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	CO (mg/m^3)	O_3 (μg/m^3)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)
BinshuiEastRoad	0.92	0.55	0.67	0.19	0.89	0.55
JiansheRoad	0.62	0.82	0.56	0.29	0.50	0.37
NorthRingRoad	0.00	0.00	0.22	1.00	0.00	0.00
Xisido	0.62	0.61	0.22	0.19	0.63	0.55
ZhongshanNorthRoad	0.31	0.74	0.44	0.00	0.85	0.53
DaliRoad	0.54	0.76	0.33	0.13	0.54	0.61
BinshuiWestRoad	0.38	0.55	0.22	0.00	0.61	0.61
JinguRoad	0.62	0.82	0.44	0.06	0.78	0.79
HexiYijingRoad	0.85	0.55	0.56	0.29	0.48	0.42
DiweiRoad	0.69	0.61	0.78	0.32	0.80	1.00
XinlaoRoad	0.77	1.00	1.00	0.16	1.00	0.97
YongyangWestRoad	0.38	0.66	0.56	0.19	0.74	0.47
Tuanpowa	0.00	0.74	0.44	0.13	0.80	0.79
HanbeiRoad	0.77	0.42	0.78	0.16	0.48	0.39
YongmingRoad	1.00	0.95	0.67	0.16	0.70	0.76
FourthStreet	0.77	0.74	1.00	0.00	0.46	0.63
LeapForwardRoad	0.31	0.84	0.56	0.19	0.93	0.74
HuaiheRoad	0.46	0.50	0.00	0.10	0.54	0.47
Forwardlane	0.69	0.79	0.44	0.19	0.83	0.71
DazhiguNo8Road	0.46	0.82	0.56	0.06	0.30	0.63
Thepathofdiligenceandfrugality	0.77	0.84	0.33	0.03	0.61	0.39
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Yanqing Summer Capital	0.45	0.42	0.38	0.57	0.31	0.19
Miyun New City	0.25	0.60	0.50	0.29	0.48	0.47
Pinggu New City	0.10	0.00	0.50	0.86	0.04	0.06
Miyun Town	0.15	0.74	0.38	0.14	0.27	0.28
Fengtai Xiaotun	0.25	0.98	0.75	0.29	0.85	0.56
Huairou New City	0.45	0.74	0.50	0.24	1.00	1.00
Yanqing Shiheying	0.15	0.81	0.50	0.38	0.67	0.53
Daxing Old Palace	1.00	0.91	0.38	0.10	0.65	0.56
Fangshan and Yanshan	0.30	0.60	0.75	0.43	0.40	0.38
Tongzhou Dongguan	0.00	0.72	0.75	0.43	0.37	0.34
Fengtai Yungang	0.25	0.67	0.75	0.52	0.37	0.28
Three stores in Mentougou	0.10	0.42	0.25	1.00	0.02	0.06
ancient city	0.15	0.86	0.13	0.38	0.63	0.47

(Continued on following page)

TABLE 9 (*Continued*) Standardized air pollutant concentrations across different regions.tianjing

Region	Pollutant concentration					
	SO_2 (μg/m³)	NO_2 (μg/m³)	CO (mg/m³)	O_3 (μg/m³)	PM_{10} (μg/m³)	$PM_{2.5}$ (μg/m³)
Olympic Sports Center	0.00	0.44	0.00	0.90	0.00	0.00
Changping Town	0.25	0.70	0.38	0.52	0.46	0.44
Huairou Town	0.50	0.98	0.25	0.00	0.37	0.28
Shunyi New City	0.15	0.79	0.50	0.38	0.71	0.53
Haidian Wanliu	0.30	1.00	1.00	0.24	0.77	0.53
Official Garden	0.25	0.81	0.75	0.52	0.50	0.44
Agricultural Exhibition Hall	0.20	0.86	0.38	0.33	0.52	0.44
Temple of Heaven	0.35	0.91	1.00	0.52	0.62	0.47

Let mA and mB be sets of BPAs corresponding to focal elements $A_1, A_2 \ldots\ldots A_i$ and $B_1, B_2 \ldots\ldots B_j$, respectively. Let m denote the new evidence after combining mA and mB. Then, the Dempster combination rule is expressed as follows:

$$m(\varnothing) = 0 \tag{8}$$

$$m(A) = \frac{1}{1-k}\sum_{(A_i \cap B_j)\neq\varnothing} m_A(A_i) \cdot m_B\left(B_j\right) \tag{9}$$

$$k = \sum_{(A_i \cap B_j)\neq\varnothing} m_A(A_i) \cdot m_B\left(B_j\right) \tag{10}$$

where K is the conflict coefficient, which reflects the degree of conflict between focal elements. A larger K indicates greater conflict; the combination rule cannot be used when $K = 1$.

2.1.3 Support vector machine (SVM)

We derived a data fusion method to assess air pollution based on comprehensive fuzzy evaluation, the Dempster–Shafer (DS) evidence theory, and the KNN algorithm. To comprehensively evaluate this method, we chose the SVM method as a benchmark for comparison. SVM, a widely used supervised learning method for classification and regression, is based on structural risk minimization. SVM can simultaneously process linear and non-linear data via kernel functions, making it suitable for analyzing our dataset.

2.2 Model overview

2.2.1 Research concept

The atmospheric environment is complex, and air quality is influenced by many factors, including uncertain and fuzzy factors (Seinfeld and Pandis, 2016). To address this environmental complexity, we propose an air pollution concentration evaluation method based on the DS evidence theory corrected by subjective and objective weighting and the extensible KNN. First, the set of air pollution evaluation indicators is determined (Figure 2). Figure 3 presents the air-pollution modeling methodology and workflow. The workflow encompasses each phase from the initial data segmentation to the final evaluation of pollution levels and elucidates the seamless transition between phases. This method thus provides a cohesive approach to quantifying air pollution.

2.2.2 Research procedure

Comprehensive examination and iterative testing have revealed that accurate quantification of air pollution requires consideration of both the pollution intensity for specific contaminants and the spatiotemporal dynamics of the atmosphere. We therefore designed and refined an optimized algorithm combining evidence theory-based data fusion and KNN-based fuzzy evaluation. This algorithm was not chosen arbitrarily; it emerged as the most effective after several experimental iterations and methodological trials. Its efficacy in addressing the complexity of pollution assessment underscores its robustness. This integrated algorithm is deployed by applying evidence theory, the KNN algorithm, and then their synthesis.

2.2.2.1 Evidence theory application

For data segmentation, the air pollution data are initially divided into different intervals, based on national air pollution concentration standards, to ensure consistency and accuracy in evaluation.

Basic Probability Assignment (BPA) values are constructed for each pollutant, based on interval similarity. This step underlies evidence theory fusion, with the key objective being the accurate representation of uncertainty in pollutant concentration.

For each pollutant, evidence theory is applied to sequentially fuse the concentration data from each region, thereby enhancing the model's capacity to handle regional disparities.

2.2.2.2 KNN algorithm application

Selecting the value of k: The optimal k value is determined based on cross-validation results to ensure the model's accuracy and generalizability.

Distance Measurement: An improved method of distance measurement is applied to calculate the similarity between data points, aiding in more accurately identifying the nearest neighbors.

2.2.2.3 Synthesis of evidence theory and KNN

The evidence theory fusion results are further corrected using subjective and objective weights to optimize assessment accuracy.

The output of the evidence theory computations is used as input for the KNN algorithm, and the concentrations are estimated by

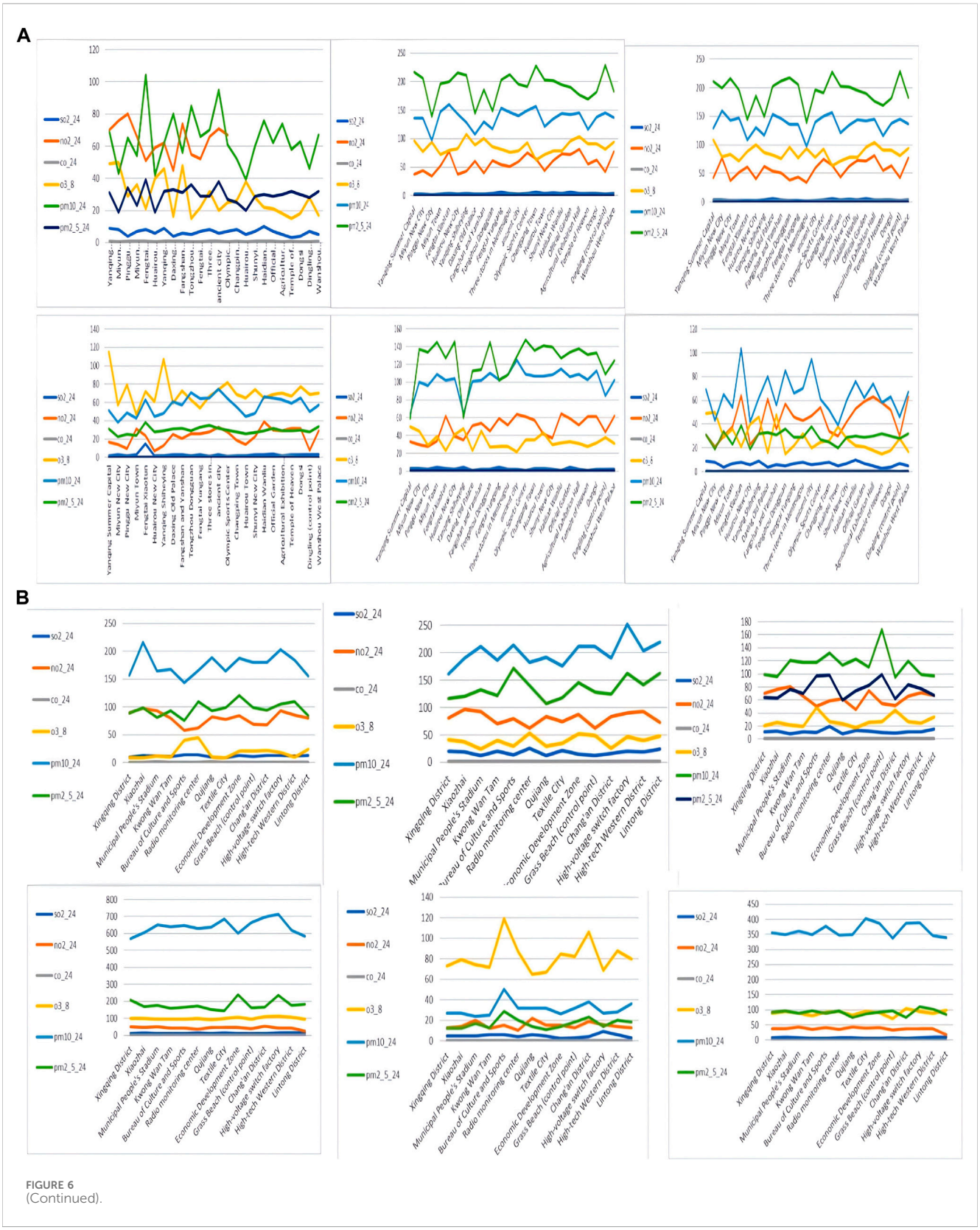


FIGURE 6
(Continued).

applying the matter-element extension model. This step accounts for the spatiotemporal dynamics of air pollutants and their environmental and health impacts.

Our model thus enhances the accuracy of air pollution assessment based on multi-source data and effectively fuses the data to improve the comprehensiveness and reliability of the results.

FIGURE 6
(Continued). Comparison of the monitored concentrations of different pollutants in different regions. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

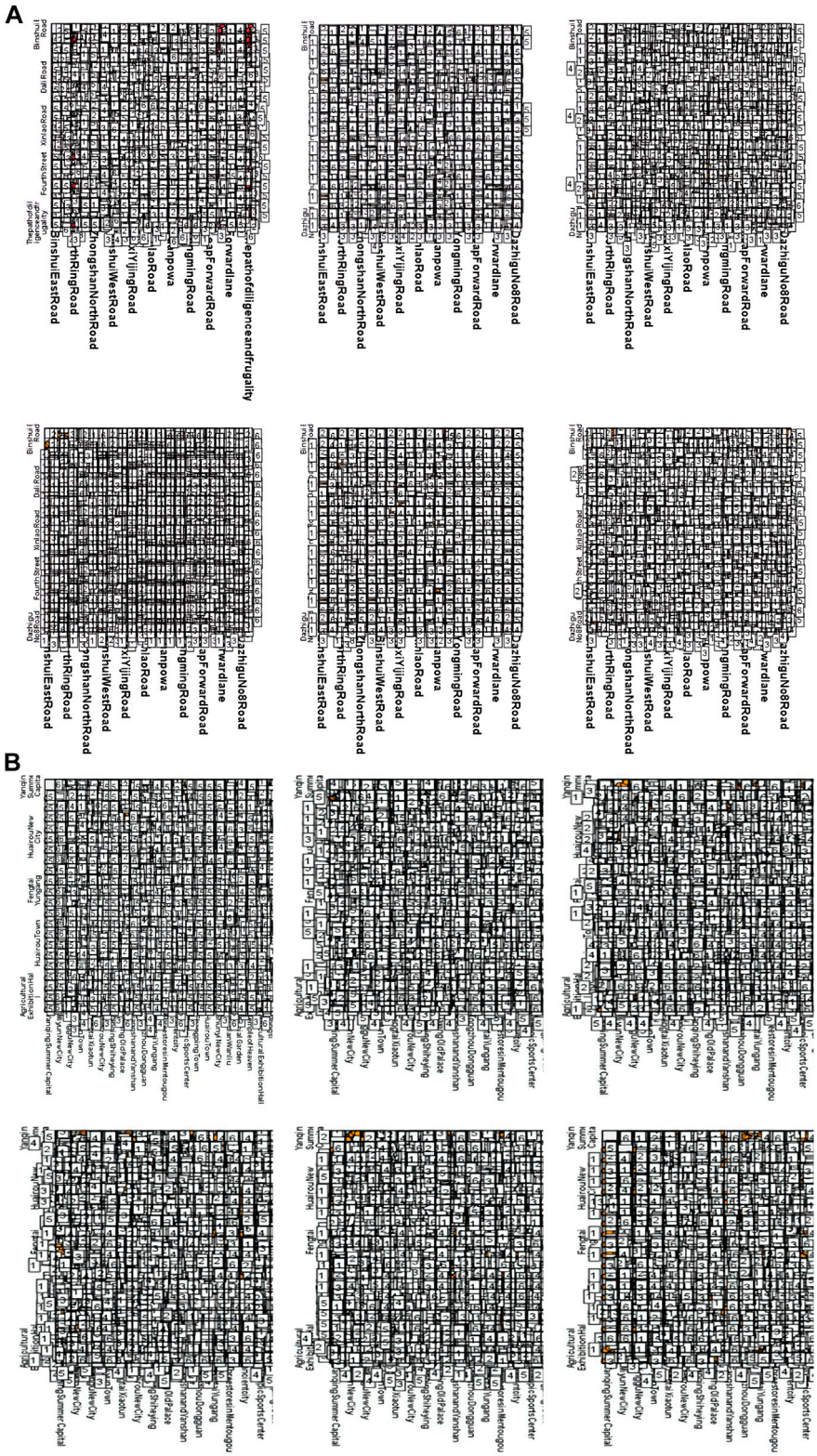


FIGURE 7
(Continued).

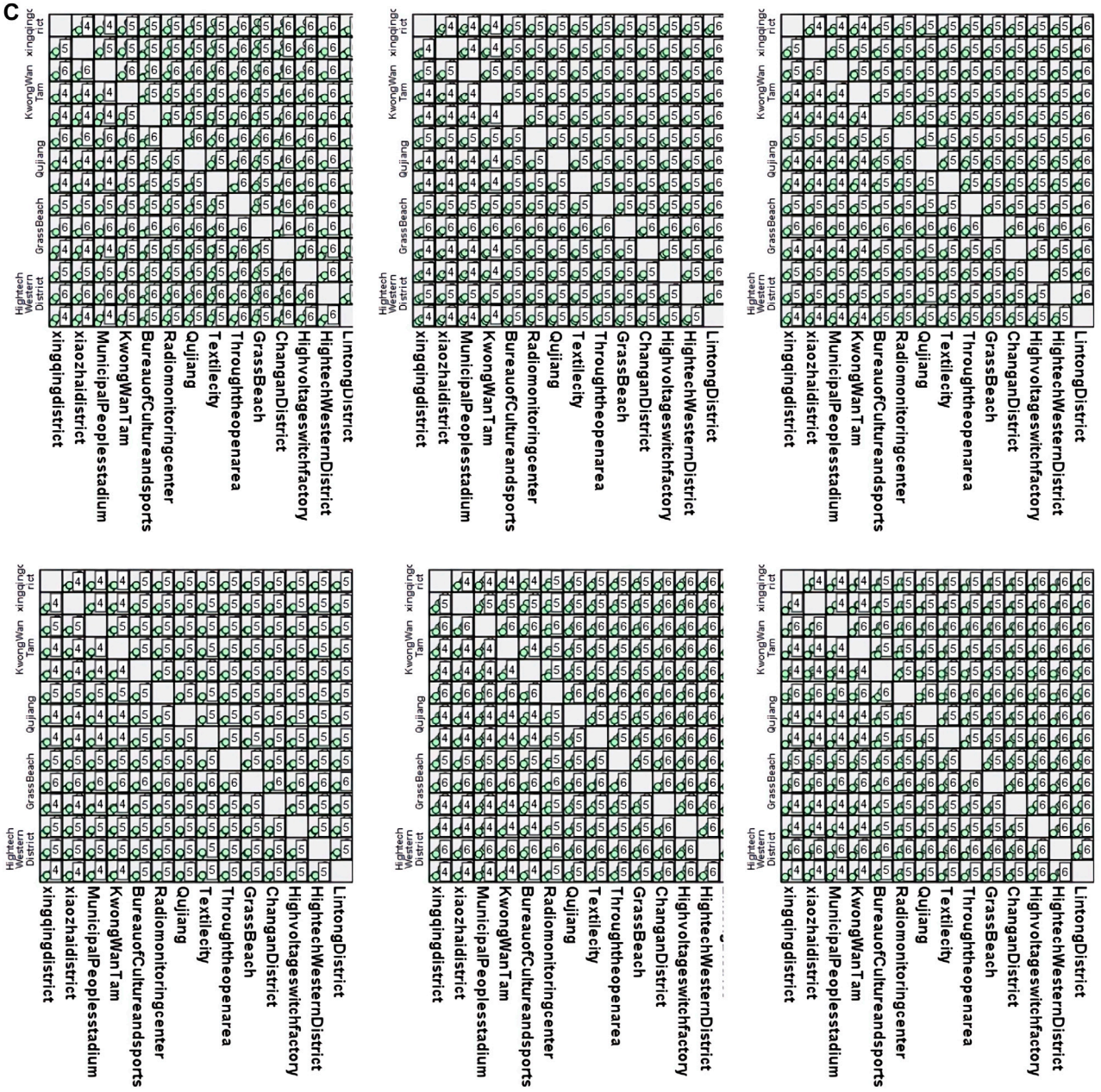


FIGURE 7
(Continued). Correlations among pollutant concentrations in different regions. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

2.3 Research methods

2.3.1 Fusion of air pollutant concentrations based on interval similarity

The air pollutant concentrations were fused as follows:

a) For each urban area within the city, a distinct set (U) of pollutant concentrations was determined:

$$U = \{U_1, U_2 \cdots U_N\} \tag{11}$$

Each single factor was then decomposed into subsystem evaluation factors, with the corresponding factor set U_i represented as

$$U_i = \{U_{i1}, U_{i2} \cdots U_{iN}\} \tag{12}$$

where U_i is the factor set and U_{iN} is the *Nth* sub-factor indicator in the *ith* factor set.

b) Based on the degree of harm to human health caused by air pollution, the air pollution concentrations were categorized into seven levels, thus determining the evaluation set (Table 2), denoted as

$$V = \{V_1, V_2 \cdots V_7\} \tag{13}$$

where V_1–V_7 correspond to seven levels of air pollution concentration.

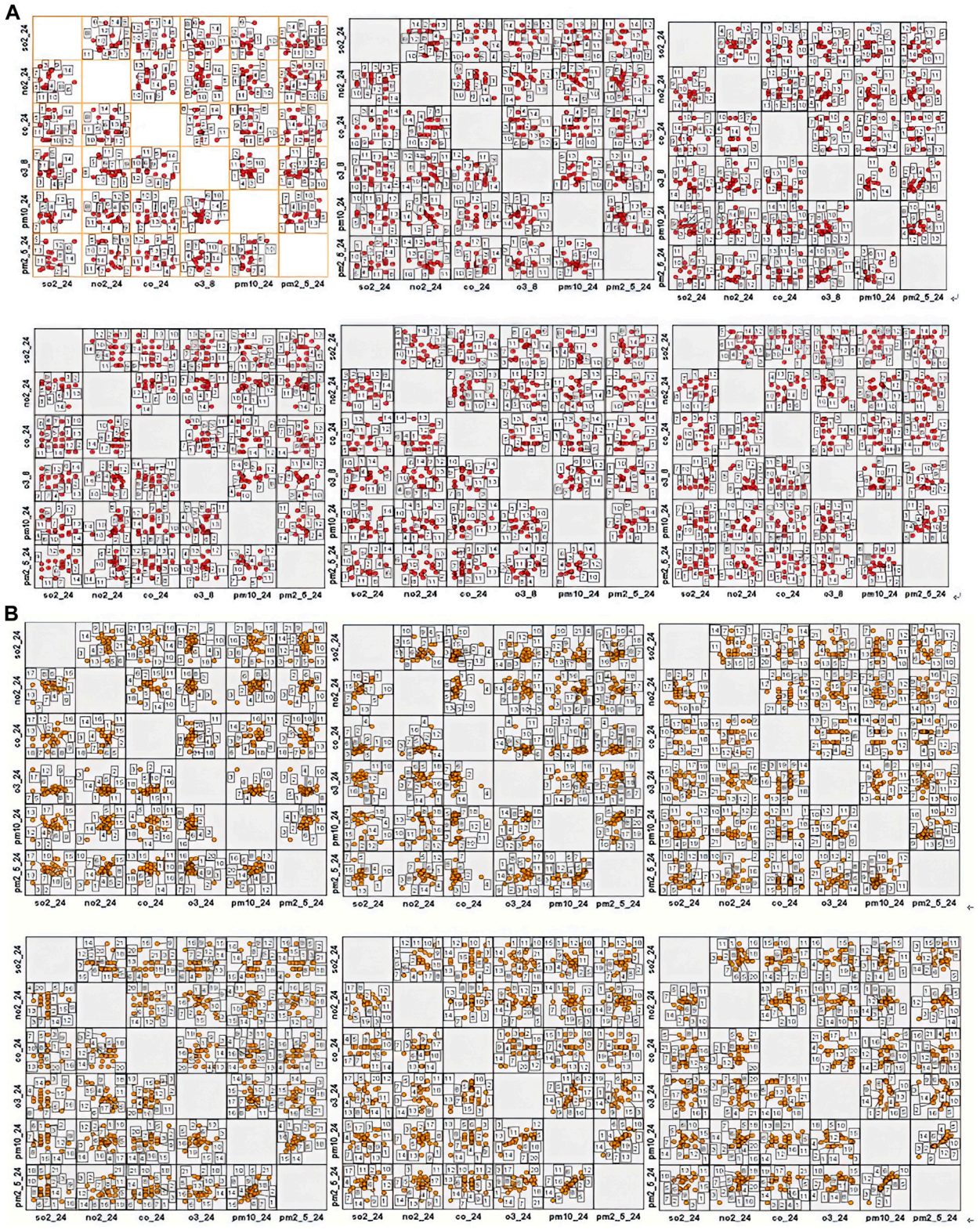


FIGURE 8
(Continued).

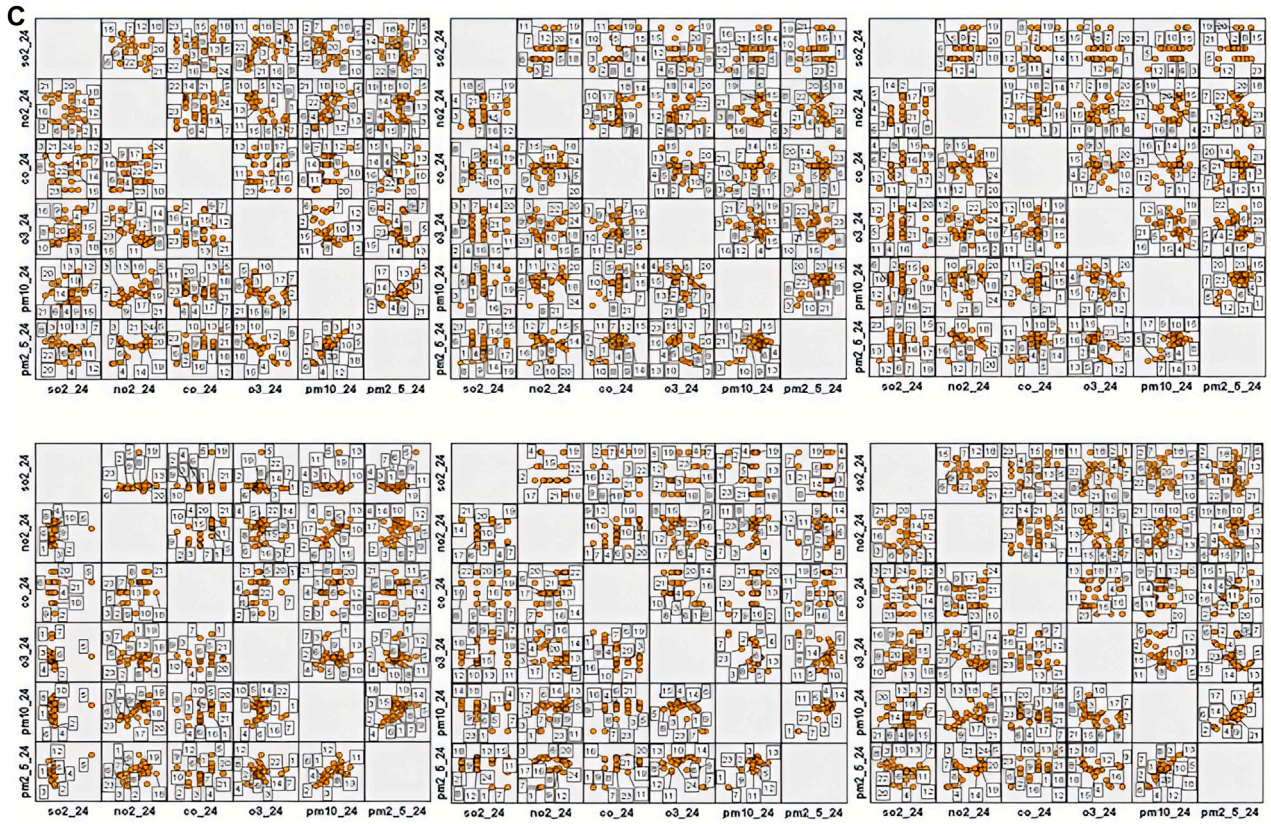


FIGURE 8
Correlations among pollutant concentrations. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

c) To determine the membership function, based on the air pollution concentration standards of the Chinese Environmental Protection Administration (EPA), the air pollutant concentrations were classified by interval. For each pollutant, the distance to each interval was calculated using Equation (14). A larger distance to the interval corresponds to a lower similarity and a greater impact of pollutants on air pollution concentrations.

$$D(U, V_k) = \frac{\sum_{k=1}^{n} |U - V_k|}{n} \tag{14}$$

Similarity was then calculated via Equation 15 and was normalized to yield the basic probability distribution function, as follows:

$$S(U, V_k) = \frac{1}{(1 + D(U, V_k))} \tag{15}$$

d) The DS evidence from different regions for each pollutant was then sequentially fused. Following the first law of geography, the law of spatial correlation, which states that "all things are related, but nearby things are more related than distant things" (Tobler, 1970; Li et al., 2016), it is necessary to consider the impact of pollutant concentrations and spatiotemporal influences. During data fusion, data transmission can be delayed. For greater efficiency, we applied sequential data fusion, fusing the output data items one-by-one, based on their to spatial sequence. Unlike other fusion techniques, sequential data fusion better balances data quality and processing speed, enabling real-time or near-real-time air quality assessments. Figure 4 illustrates BPA-value adjustment using interval similarity via DS evidence theory. This adjustment effectively addresses the inherent uncertainty associated with atmospheric pollution concentrations. Merging the modified BPA values produces the final atmospheric pollution concentrations:

$$U_N = (U_1, U_2, U_3 \cdots U_N) \tag{16}$$

2.3.2 Construction of the DS evidence theory model based on comprehensive weighting

To evaluate the air pollution concentrations, concentrations of air pollutants from multiple sources were fused. This analysis requires consideration of individual pollutant concentrations and of the mutual influences among the pollution levels in different regions for each pollutant. Therefore, each indicator has both subjective and objective weights. Subjective (or expert) weighting is derived from expert opinion or experience. Objective weighting is based on the relationships among indicators and relies on specific mathematical methods to calculate the weight of each indicator.

TABLE 10 BPA values across different regions.xi'an.

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Xingqing District	0.1334	0.723,083	0.72041	0.7047	0.68038	0.6795
Xiaozhai	0.1335	0.73219	0.71237	0.7116	0.67454	0.6764
Municipal People's Stadium	0.133	0.738,389	0.71503	0.706	0.72652	0.7176
Kwong Wan Tam	0.1334	0.717,136	0.71771	0.7033	0.719,877	0.6987
Bureau of Culture and Sports	0.1332	0.695,682	0.71771	0.7437	0.719,877	0.7641
Radio monitoring center	0.1345	0.706,962	0.71237	0.713	0.75202	0.7665
Qujiang	0.133	0.711,287	0.71237	0.7088	0.71118	0.6673
Textile City	0.1336	0.687,456	0.72041	0.7006	0.73103	0.7121
Economic Development Zone	0.1335	0.729,129	0.71237	0.7116	0.70269	0.7325
Grass Beach (control point)	0.1332	0.701,277	0.70972	0.713	0.82813	0.7689
Chang'an District	0.1331	0.697,073	0.70972	0.7377	0.67262	0.6733
High-voltage switch factory	0.1334	0.717,136	0.70972	0.713	0.72429	0.7346
High-tech Western District	0.1334	0.724,585	0.71237	0.7102	0.68038	0.7197
Lintong District	0.1338	0.718,614	0.71771	0.7229	0.7265	0.676,476
Xingqing District	0.13472	0.71909	0.71463	0.7074	0.6335	0.671,995
Xiaozhai	0.13508	0.7267	0.72273	0.7102	0.7450	0.681,060
Municipal People's Stadium	0.13484	0.72855	0.72273	0.7189	0.6805	0.68106
Kwong Wan Tam	0.13472	0.72060	0.70,670,944	0.68828	0.654,218	0.69672
Bureau of Culture and Sports	0.13484	0.690,182	0.709,331	0.7524	0.64844	0.68413
Radio monitoring center	0.135,089	0.69437	0.714,633	0.69911	0.8218	0.82859
Qujiang	0.13472	0.71312	0.714,633	0.69365	0.814,504	0.73031
Textile City	0.13447	0.713,121	0.71,197,217	0.6923	0.66209	0.68724
Economic Development Zone	0.13508	0.73041	0.714,633	0.71601	0.7717	0.726,812,618
Grass Beach (control point)	0.134,478	0.708,705	0.71,197,217	0.70187	0.742,507	0.752,074,228
Chang'an District	0.134,844	0.690,182	0.70,670,944	0.73680	0.69557	0.663,168,781
High-voltage switch factory	0.135,212	0.726,707	0.717,313	0.73833	0.75421	0.741,035,957
High-tech Western District	0.135,089	0.722,122	0.72,001,489	0.73987	0.68475	0.723,341,689
Lintong District	0.135,212	0.716,096	0.71,197,217	0.704,659	0.6912	0.732,435,926
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Xingqing District	0.1,347,224	0.719,096	0.714,633	0.707,465	0.633,503	0.67,199,534
Xiaozhai	0.13,508,933	0.726,707	0.722,736	0.71029	0.74502	0.681,060,025
Municipal People's Stadium	0.1,348,445	0.728,558	0.7,227,363	0.718,915	0.680,516	0.681,060,025
Kwong Wan Tam	0.1,347,224	0.720,606	0.70,670,944	0.68828	0.65422	0.696,723,779
Bureau of Culture and Sports	0.1,348,445	0.6,901,826	0.709,331	0.752,404	0.648,435	0.684,136,177
Radio monitoring center	0.1,350,893	0.69437	0.714,633	0.69911	0.821,803	0.828,598,193

(Continued on following page)

TABLE 10 (*Continued*) BPA values across different regions.xi'an.

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Qujiang	0.1,347,224	0.7,131,217	0.71,463,303	0.69365	0.814,504	0.730,317,018
Textile City	0.1,344,789	0.7,131,217	0.71,197,217	0.6923	0.66209	0.687,240,243
Economic Development Zone	0.1,350,893	0.730,418	0.714,633	0.716,018	0.771,665	0.726,812,618
Grass Beach (control point)	0.1,344,789	0.708,705	0.71,197,217	0.70187	0.742,507	0.752,074,228
Chang'an District	0.1,348,445	0.690,182	0.70,670,944	0.7368	0.695,572	0.663,168,781
High-voltage switch factory	0.135,212	0.726,707	0.7,173,138	0.73833	0.754,212	0.741,035,957
High-tech Western District	0.13,508,933	0.722,122	0.72,001,489	0.73987	0.684,751	0.723,341,689
Lintong District	0.13,521,208	0.7,160,968	0.71,197,217	0.704,659	0.6,912,027	0.732,435,926

2.3.3 Subjective weights

Each air pollutant has a specific toxicity level, with higher toxicity correlating with increased pollution and potential harm to humans. To ensure a well-rounded and credible representation of pollutant toxicity, we derived subjective weights (denoted by W_i) for each pollutant, by consulting a panel of experts and corroborating their input using the literature. This approach both leverages expert experience and ensures a balanced weighting strategy supported by established knowledge.

2.3.4 Objective weights

Objective weighting has the advantage of relying primarily on objective data without requiring human intervention. Therefore, objective judgments are not limited by human subjectivity and are considered true and reliable.

Air pollution levels depend both on differences between pollutants and regional pollutant concentrations. Therefore, to calculate the objective weights, we implemented the CRIteria Importance Through Intercriteria Correlation (CRITIC) method (Wang and Jiang, 2017; Ying and Yunyun, 2017), a sophisticated approach that effectively integrates both contrasting and conflicting indicators. In this context, associations between information are used to quantify the conflict between indicators. More closely associated indicators exhibit less conflict, and are therefore assigned less weight. The contrast between indicators was ascertained using the standard deviation of the information; a larger standard deviation implies a greater contrast. Here, we adopted the "Over-standard Multiple Method" to represent this contrast, using the standard deviation of information. Indicator similarity, conversely, was used to compute the conflict between indicators. Among the various weight-calculation methods, CRITIC was deemed the most suitable for our research objectives; the other methods, in contrast, exhibit greater computational complexity, specific data prerequisites, or inherent limitations. The objective weights were calculated as follows.

a) According to Chinese EPA monitoring data, the regions $P_1, P_2, P_3 I ... P_j$ are currently monitored for the pollutants SO_2, NO_2, CO, O_3, $PM_{2.5}$, and PM_{10}. The daily air pollutant concentrations in the different regions are as follows:

$$U = \begin{Bmatrix} U_{11}\,U_{12}\,U_{13} \ldots\ldots\ldots\ldots\ldots U_{1N} \\ U_{21}\,U_{22}\,U_{23} \ldots\ldots\ldots\ldots\ldots U_{2N} \\ U_{31}\,U_{32}\,U_{33} \ldots\ldots\ldots\ldots\ldots U_{3N} \\ U_{41}\,U_{42}\,U_{43} \ldots\ldots\ldots\ldots\ldots U_{4N} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ U_{i1}\,U_{I2}\,U_{i3} \ldots\ldots \ldots\ldots \ldots U_{iN} \end{Bmatrix} \tag{17}$$

where N is the number of air pollutants (SO_2, NO_2, CO, O_3, $PM_{2.5}$, and PM_{10}) and i is the monitoring region.

b) To better demonstrate pollution levels, the pollutant concentration indicators for the different regions and pollutants in the matrix were nondimensionalized. The concentrations of the same pollutant in different regions mutually affect one another. Therefore, range normalization was used for nondimensionalization:

$$U_{iN}/ = \frac{U_{iN} - \min}{\max - \min} \tag{18}$$

After nondimensionalization, the relative pollution levels for the same pollutant in different regions are obtained, as follows:

$$\mathrm{U} = \begin{Bmatrix} U_{11}\,U_{12}\,U_{13} \ldots\ldots\ldots\ldots\ldots U_{1N} \\ U_{21}\,U_{22}\,U_{23} \ldots\ldots\ldots\ldots\ldots U_{2N} \\ U_{31}\,U_{32}\,U_{33} \ldots\ldots\ldots\ldots\ldots U_{3N} \\ U_{41}\,U_{42}\,U_{43} \ldots\ldots\ldots\ldots\ldots U_{4N} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ U_{i1}\,U_{I2}\,U_{i3} \ldots\ldots \ldots\ldots \ldots U_{iN} \end{Bmatrix} \tag{19}$$

where N is the number of air pollutants (SO_2, NO_2, CO, O_3, $PM_{2.5}$, and PM_{10}) and i is the monitoring region.

c) The concentrations of different pollutants in the same region can also mutually affect one another. When multiple different pollutants are present in greater concentrations, their impact on one another is greater. The contrast among indicators in the CRITIC method can be represented by over-standard multiples, as follows:

TABLE 11 BPA values across different regions.beijing

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Yanqing Summer Capital	0.41,729,175	0.41	0.41737	0.414,353	0.408,899	0.408,415
Miyun New City	0.41,691,635	0.42	0.41893	0.414,353	0.415,314	0.422,296
Pinggu New City	0.41,691,635	0.42	0.415,822	0.418,275	0.413,154	0.413,356
Miyun Town	0.41,654,162	0.41	0.414,285	0.426,345	0.405,764	0.39772
Fengtai Xiaotun	0.41,654,162	0.41	0.415,822	0.415,131	0.406,804	0.415,871
Huairou New City	0.41,654,162	0.42	0.414,285	0.420,663	0.413,154	0.419,702
Yanqing Shiheying	0.41,691,635	0.42	0.41737	0.414,353	0.417,498	0.42758
Daxing Old Palace	0.41,691,635	0.42	0.41737	0.415,131	0.419,705	0.413,356
Fangshan and Yanshan	0.41,654,162	0.42	0.41893	0.415,913	0.42993	0.415,871
Tongzhou Dongguan	0.41,654,162	0.42	0.415,822	0.427,169	0.412,081	0.400,048
Fengtai Yungang	0.41,691,635	0.41	0.414,285	0.415,913	0.412,081	0.408,415
Three stores in Mentougou	0.41,654,162	0.4	0.414,285	0.423,891	0.407,849	0.405,989
ancient city	0.41,616,756	0.43	0.41737	0.412,805	0.425,325	0.423,604
Olympic Sports Center	0.41,616,756	0.42	0.41737	0.416,697	0.425,325	0.423,604
Changping Town	0.41,654,162	0.41	0.41737	0.413,577	0.427,615	0.420,995
Huairou Town	0.41,691,635	0.42	0.414,285	0.416,697	0.412,081	0.41461
Shunyi New City	0.41,691,635	0.42	0.415,822	0.415,131	0.411,015	0.410,871
Haidian Wanliu	0.41,691,635	0.42	0.41737	0.413,577	0.42993	0.420,995
Official Garden	0.41,654,162	0.43	0.420,501	0.415,131	0.423,059	0.423,604
Agricultural Exhibition Hall	0.41,654,162	0.42	0.414,285	0.415,913	0.423,059	0.428,922
Temple of Heaven	0.41,616,756	0.41	0.41893	0.413,577	0.407,849	0.423,604
Dongsi	0.41,691,635	0.42	0.41893	0.415,913	0.418,598	0.422,296
Dingling (control point)	0.41,616,756	0.42	0.415,822	0.415,913	0.415,314	0.413,356
Wanshou West Palace	0.41,691,635	0.42	0.41737	0.413,577	0.418,598	0.424,921
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Yanqing Summer Capital	0.41,729,175	0.41	0.41737	0.414,353	0.408,899	0.408,415
Miyun New City	0.41,691,635	0.42	0.41893	0.414,353	0.415,314	0.422,296
Pinggu New City	0.41,691,635	0.42	0.415,822	0.418,275	0.413,154	0.413,356
Miyun Town	0.41,654,162	0.41	0.414,285	0.426,345	0.405,764	0.39772
Fengtai Xiaotun	0.41,654,162	0.41	0.415,822	0.415,131	0.406,804	0.415,871
Huairou New City	0.41,654,162	0.42	0.414,285	0.420,663	0.413,154	0.419,702
Yanqing Shiheying	0.41,691,635	0.42	0.41737	0.414,353	0.417,498	0.42758
Daxing Old Palace	0.41,691,635	0.42	0.41737	0.415,131	0.419,705	0.413,356
Fangshan and Yanshan	0.41,654,162	0.42	0.41893	0.415,913	0.42993	0.415,871
Tongzhou Dongguan	0.41,654,162	0.42	0.415,822	0.427,169	0.412,081	0.400,048

(Continued on following page)

TABLE 11 (*Continued*) BPA values across different regions.beijing

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Fengtai Yungang	0.41,691,635	0.41	0.414,285	0.415,913	0.412,081	0.408,415
Three stores in Mentougou	0.41,654,162	0.4	0.414,285	0.423,891	0.407,849	0.405,989
ancient city	0.41,616,756	0.43	0.41737	0.412,805	0.425,325	0.423,604
Olympic Sports Center	0.41,616,756	0.42	0.41737	0.416,697	0.425,325	0.423,604
Changping Town	0.41,654,162	0.41	0.41737	0.413,577	0.427,615	0.420,995
Huairou Town	0.41,691,635	0.42	0.414,285	0.416,697	0.412,081	0.41461
Shunyi New City	0.41,691,635	0.42	0.415,822	0.415,131	0.411,015	0.410,871
Haidian Wanliu	0.41,691,635	0.42	0.41737	0.413,577	0.42993	0.420,995
Official Garden	0.41,654,162	0.43	0.420,501	0.415,131	0.423,059	0.423,604
Agricultural Exhibition Hall	0.41,654,162	0.42	0.414,285	0.415,913	0.423,059	0.428,922
Temple of Heaven	0.41,616,756	0.41	0.41893	0.413,577	0.407,849	0.423,604
Dongsi	0.41,691,635	0.42	0.41893	0.415,913	0.418,598	0.422,296
Dingling (control point)	0.41,616,756	0.42	0.415,822	0.415,913	0.415,314	0.413,356
Wanshou West Palace	0.41,691,635	0.42	0.41737	0.413,577	0.418,598	0.424,921

$$S_j = U_{iN} \Big/ \left(\frac{1}{n} \sum_{j=1}^{n} V_{ij} \right) \tag{20}$$

d) For the same pollutant, the levels are strongly positively correlated among regions. The extent of correlation can be determined from the similarity among indicators. Greater similarity corresponds to less conflict. Here, the indicator similarity matrix was constructed using Pearson correlation coefficients. The level of conflict (q_{ij}) of each indicator in the similarity matrix was calculated as follows:

$$Q_j = 1 - |q_{ij}| \tag{21}$$

e) F_j, the degree of pollution caused by the *jth* pollutant, was calculated as

$$F_j = Q_j * S_j \tag{22}$$

f) W_j, the objective weight of the *jth* pollutant, was calculated as

$$W_j = \frac{F_j}{\sum_{j=1}^{n} F_j} \tag{23}$$

2.3.5 Combined subjective and objective weights

The combined subjective and objective weights of the *jth* indicator were determined using the normalized combined weights, as follows:

$$W_j^* = \mathrm{w}_i w_j \Big/ \sum_{j=1}^{N} \mathrm{w}_i w_j \text{ where } j = 1.....N \tag{24}$$

W_i and W_j are, respectively, the subjective and objective weight elements of the *nth* indicator.

2.3.6 Subjective and objective correction

Subjective and objective correction of the concentrations of different pollutants in the same region was performed as follows:

$$D = W_j^* * \mathrm{U_N} \tag{25}$$

2.4 Construction of the corrected KNN model based on the average extensible distance

The KNN algorithm determines the class of a sample by counting the number of nearest neighbors of the same class. However, if the distribution of classified samples is highly scattered and the data distribution is not considered, the results can easily be biased. Yang et al. (2010) proposed using the extensible distance to vote and select the target attribute of the dataset for classification. Tan et al. (2017) developed an air quality evaluation model based on fuzzy matter-element analysis. Xiao and Duan (2013) proposed improving the importance of classes using attribute values. Dai et al. (2018) introduced a method to calculate sample-attribute weights using the analytic hierarchy process and to classify samples based on their weighted distances. Lü et al. (2021) proposed correcting KNN classification using the probability that the test sample belongs to each fault type as the weight.

Pollutant concentration data is typically log-normally distributed, as substantiated by Ott (1990), who provided the physical explanation for this phenomenon. This log-normal distribution is further supported by the method-fusion approach for enhancing long-term air pollution estimates, as discussed by Chastko and Adams (2019). Air pollution concentration data follow a normal distribution (Figure 5). Each

TABLE 12 BPA values across different regions.tianjin

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Xingqing District	0.385,551	0.374,772	0.386,021	0.399,306	0.392,035	0.390,947
Xiaozhai	0.385,203	0.365,038	0.381,788	0.400,129	0.364,171	0.366,603
Municipal People's Stadium	0.383,815	0.373,775	0.386,021	0.383,532	0.387,913	0.397,547
Kwong Wan Tam	0.384,855	0.381,901	0.3846	0.388,909	0.377,011	0.374,374
Bureau of Culture and Sports	0.385,203	0.402,857	0.388,896	0.377,566	0.432,226	0.406,366
Radio monitoring center	0.384,508	0.366,944	0.381,788	0.39205	0.362,917	0.366,603
Qujiang	0.385,551	0.375,774	0.387,453	0.396,858	0.384,878	0.393,122
Textile City	0.383,815	0.400,398	0.3846	0.373,931	0.403,837	0.395,322
Economic Development Zone	0.384,508	0.379,837	0.386,021	0.398,487	0.378,947	0.390,947
Grass Beach (control point)	0.384,161	0.397,165	0.383,189	0.373,212	0.40944	0.401,419
Chang'an District	0.384,855	0.388,537	0.383,189	0.378,302	0.388,935	0.386,668
High-voltage switch factory	0.385,203	0.386,249	0.387,453	0.385,818	0.39308	0.386,668
High-tech Western District	0.384,508	0.389,306	0.386,021	0.376,833	0.421,125	0.404,704
Lintong District	0.384,855	0.394,774	0.383,189	0.380,526	0.383,876	0.382,481
Xingqing District	0.385,203	0.377,794	0.3846	0.381,273	0.375,094	0.378,384
Xiaozhai	0.384,161	0.373,775	0.383,189	0.390,473	0.359,207	0.368,516
Municipal People's Stadium	0.384,855	0.385,493	0.386,021	0.382,776	0.38288	0.386,668
Kwong Wan Tam	0.3859	0.393,984	0.387,453	0.378,302	0.399,464	0.388,796
Bureau of Culture and Sports	0.384,855	0.398,775	0.386,021	0.377,566	0.384,878	0.386,668
Radio monitoring center	0.384,161	0.402,034	0.383,189	0.375,376	0.397,313	0.388,796
Qujiang	0.38347	0.397,968	0.387,453	0.373,212	0.380,904	0.393,122
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
BinshuiEastRoad	0.477,904	0.47468	0.479,314	0.477,222	0.480,541,073	0.512,594,485
JiansheRoad	0.478,341	0.478,999	0.4825	0.478,138	0.481,408,476	0.507,314,542
NorthRingRoad	0.476,164	0.475,906	0.47204	0.474,497	0.471,201,936	0.443,376,925
Xisido	0.475,299	0.479,622	0.479,314	0.467,378	0.462,218,883	0.469,810,367
ZhongshanNorthRoad	0.474,437	0.476,521	0.483,841	0.470,023	0.463,826,601	0.528,169,295
DaliRoad	0.474,867	0.475,906	0.47204	0.483,705	0.490,257,897	0.470,933,419
BinshuiWestRoad	0.473,149	0.471,642	0.473,837	0.483,705	0.491,160,766	0.47,547,984
JinguRoad	0.478,341	0.479,622	0.479,314	0.479,979	0.490,257,897	0.503,425,418
HexiYijingRoad	0.476,598	0.475,906	0.475,649	0.477,222	0.476,250,528	0.441,390,173
DiweiRoad	0.478,341	0.475,906	0.477,474	0.472,697	0.470,370,892	0.436,500,326
XinlaoRoad	0.474,867	0.465,094	0.475,649	0.470,911	0.464,634,662	0.436,500,326
YongyangWestRoad	0.476,164	0.444,523	0.470,256	0.482,768	0.479,676,791	0.352,427,812
Tuanpowa	0.475,731	0.477,138	0.47204	0.47631	0.475,401,597	0.509,940,846

(Continued on following page)

TABLE 12 (*Continued*) BPA values across different regions.tianjin

Region	Pollutant concentration					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
HanbeiRoad	0.474,437	0.468,048	0.470,256	0.473,595	0.470,370,892	0.436,500,326
YongmingRoad	0.480,097	0.476,521	0.473,837	0.482,768	0.48,846,208	0.480,114,899
FourthStreet	0.477,468	0.483,397	0.477,474	0.470,023	0.473,712,781	0.486,037,369
LeapForwardRoad	0.476,164	0.485,307	0.479,314	0.477,222	0.475,401,597	0.504,284,506
HuaiheRoad	0.475,299	0.487,232	0.477,474	0.470,911	0.464,634,662	0.507,314,542
Forwardlane	0.475,299	0.47468	0.47204	0.482,768	0.484,909,629	0.480,114,899
DazhiguNo8Road	0.475,731	0.480,247	0.473,837	0.478,138	0.482,279,016	0.504,284,506
Thepathofdiligenceandfrugality	0.475,299	0.4931	0.4825	0.470,023	0.463,021,347	0.513,485,179

TABLE 13 Sequential fusion results.xi'an.

Method	Pollutant					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
D-S Sequential Fusion Results1	9.21366E-13	0.01426	0.01430	0.01428	0.01401	0.01408
		7,936	647	124	5,764	2,376
D-S Sequential Fusion Results2	5.8882	0.01415	0.0141	0.01415	0.01394	0.01416
	6E-13	5,479	85,013	2,786	6,904	668
D-S Sequential Fusion Results2	1.001	0.02054	0.02057	0.0205	0.03142	0.02046
	67E-12	5,705	6,677	60,332	1758	5,258
D-S Sequential Fusion Results2	4.3349	0.01447	0.01450	0.0144	0.0184	0.0144
	5E-13	6,865	1,514	7,182	79,643	17,647
D-S Sequential Fusion Results2	6.015	0.01407	0.01416	0.0137	0.01102	0.01296
	14E-13	9,683	2,273	7,331	1,291	1,657
D-S Sequential Fusion Results2	6.2820	0.01423	0.01427	0.01422	0.01410	0.0141
	2E-13	8,819	6,197	9,081	1,476	6,153

TABLE 14 Sequential fusion results.beijing

Method	Pollutant					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
D-S Sequential Fusion Results1	1.14978E-10	1.35918E-10	1.25262E-10	1.1126E-10	1.44329E-10	1.4529E-10
D-S Sequential Fusion Results2	7.94276E-10	8.48062E-10	8.56334E-10	8.5093E-10	8.35209E-10	8.5155E-10
D-S Sequential Fusion Results2	7.92718E-10	8.45955E-10	8.52807E-10	8.5156E-10	8.36499E-10	8.4969E-10
D-S Sequential Fusion Results2	7.95042E-10	8.52818E-10	8.57729E-10	8.4851E-10	8.50694E-10	8.5444E-10
D-S Sequential Fusion Results2	7.94844E-10	8.5099E-10	8.57506E-10	8.5514E-10	8.43495E-10	8.2224E-10
D-S Sequential Fusion Results2	7.9371	8.5339	8.54802	8.5412	8.5186	8.510
	2E-10	5E-10	E−10	E−10	6E-10	3E-10

TABLE 15 Sequential fusion results.tianjin

Method	Pollutant					
	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
D-S Sequential Fusion Results1	1.8865E-07	2.07734E-07	2.08179E-07	2.0799E-07	2.06534E-07	2.0539E-07
D-S Sequential Fusion Results2	1.88537E-07	2.06473E-07	2.06968E-07	2.0539E-07	1.90604E-07	2.074E-07
D-S Sequential Fusion Results2	1.89902E-07	2.10487E-07	2.10577E-07	2.1044E-07	2.08663E-07	2.0787E-07
D-S Sequential Fusion Results2	1.89337E-07	2.08476E-07	2.09041E-07	2.0859E-07	2.07962E-07	2.0802E-07
D-S Sequential Fusion Results2	1.89327E-07	2.08873E-07	2.09097E-07	2.0851E-07	2.05306E-07	2.0632E-07
D-S Sequential Fusion Results2	1.89425E-07	2.08922E-07	2.09619E-07	2.0953E-07	1.93548E-07	2.0614E-07

pollutant concentration has a standard interval, according to the Chinese EPA standards. Given these standard intervals, we can determine the air pollution concentrations via DS evidence theory fusion and corrected via subjective and objective weighting. Here, we propose a model algorithm based on KNN corrected using matter-element average extensible distances (Ott, 1990. A Physical Explanation of the Lognormality of Pollutant).

The matter-element extension theory (Xiang, 2008) considers the objective world as comprising matter elements; contradictions in the objective world are treated as contradictions among matter elements. Matter elements have many types of features, each with a value. Here, the matter elements (the evaluation subjects, denoted as U) and their features (the evaluation indicators, denoted as T) are combined with the values Y (which correspond to the features), forming R = (U, T, Y), the three matter-elements. The model calculations are as follows.

(1) If the described matter U has n features T_1, T_2, ... , T_m with corresponding values Y_1, Y_2, ... , Y_m, then R is called an n-dimensional fuzzy matter-element (Zhang, 1997). Combining n-dimensional matter-elements of m matters forms n-dimensional composite fuzzy matter-elements R_{nm} of m matters:

$$R = \begin{bmatrix} U & T_1 & Y_1 \\ & T_2 & Y_2 \\ & \vdots & \vdots \\ & Tn & Yn \end{bmatrix} \tag{26}$$

$$R = \begin{bmatrix} & T_1 & T_2 & \cdots & T_n \\ U_1 & Y_{11} & Y_{12} & \cdots & Y_{1n} \\ U_2 & Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ U_m & Y_{m1} & Y_{m2} & \cdots & Ymn \end{bmatrix} \tag{27}$$

(2) The described matter U has a value of Y for each feature. Let us assume that Y is within the range (a, b) belonging to a classified sample; for $X(X_1, I.....X_i)$, the samples to be classified, the aim is to determine the feature interval to which they belong, using the KNN algorithm. Therefore, we set $K = 3$ and use the quartile method to divide the interval Y of each known classified sample into three adjacent intervals.

(3) The average distance between the sample to be classified and the three nearest neighbor intervals of different classes is calculated using the interpoint distance:

$$d(x_i, y_n) = \left| x - \frac{(a_m + b_m)}{2} \right| - \frac{(b_m - a_m)}{2} \tag{28}$$

(4) The average distance between the sample to be classified and the nearest neighbor points of each class is calculated as follow:

$$\bar{d}(x_i, y_n) = \frac{1}{3} \sum \left(\left| x - \frac{(a_m + b_m)}{2} \right| - \frac{(b_m - a_m)}{2} \right) \tag{29}$$

(5) The average distance between the sample to be classified and each classified sample point is then compared. A smaller distance indicates that the sample to be classified more likely belongs to the corresponding interval.

3 Results

3.1 Evaluation indicator set

In evaluating air pollution, pollutant concentrations were calculated based on regional data for each city. It is assumed that each city has i monitoring regions, each with N indicators $u_1, u_2 \ldots .u_N$, which form the set $u = \{u_1, u_2 \ldots .u_N\}$. Each indicator set can be decomposed into subsystem evaluation factors, and the corresponding factor set U_i is represented as $u_i = \{u_{i1}, u_{i2} \ldots .u_{iN}\}$, i.e., the evaluation factor set for each region. Based on the Chinese EPA's air pollution concentration standards, an evaluation set {I, II, III, IV, V, VI, VII} is defined, where a higher level indicates a higher pollutant concentration and a greater impact on health (Table 2). The data are then nondimensionalized (Table 3).

3.2 Data source

The air quality data used here—the daily air quality data for 2022 for Xi'an, Beijing, and Tianjin—were obtained primarily from the China National Environmental Monitoring Network. Owing to space constraints, only a subset of the data is presented (Tables 4–6).

Additionally, we based our in-depth statistical analyses on the processed and analyzed datasets of the Qingyue Open Environmental Data Center. We standardized the concentrations of the various air pollutants, obtaining

TABLE 16 Subjective and objective weights of the pollutants.xi'an.

Pollutant	Indicator item				
	Indicator variability	Indicator conflict	Objective weight (%)	Subjective weight	Weight
SO_{2_24}	3.082	4.687	5.11	2	0.2511
NO_{2_24}	10.324	6.23	22.75	2	0.6275
$CO_{\_24}$	0.144	5.61	0.29	1	0.2029
O_{3_8}	8.666	5.632	17.26	3	0.7726
PM_{10_24}	19.343	4.972	34.02	4	1.1402
$PM_{2.5\_24}$	13.719	4.239	20.57	5	1.2057
SO_{2_24}	4.312	4.64	7.35	2	0.2735
NO_{2_24}	11.006	5.212	21.09	2	0.6109
$CO_{\_24}$	0.15	3.612	0.20	1	0.202
O_{3_8}	9.725	4.344	15.53	3	0.7553
PM_{10_24}	22.635	3.7	30.78	4	1.1078
$PM_{2.5\_24}$	19.773	3.447	25.05	5	1.2505
SO_{2_24}	1.598	4.488	1.77	2	0.2177
NO_{2_24}	6.796	4.858	8.15	2	0.4815
$CO_{\_24}$	0.107	6.278	0.17	1	0.2017
O_{3_8}	6.188	3.931	6.01	3	0.6601
PM_{10_24}	42.135	4.62	48.06	4	1.2806
$PM_{2.5\_24}$	28.384	5.114	35.84	5	1.3584
SO_{2_24}	1.626	6.091	4.02	2	0.2402
NO_{2_24}	6.627	4.852	13.06	2	0.5306
$CO_{\_24}$	0.141	5.893	0.34	1	0.2034
O_{3_8}	21.658	4.886	42.97	3	1.0297
PM_{10_24}	8.483	5.672	19.54	4	0.9954
$PM_{2.5\_24}$	9.913	4.988	20.08	5	1.2008
SO_{2_24}	1.73	4.985	7.68	2	0.2768
NO_{2_24}	0.109	3.662	0.35	2	0.4035
$CO_{\_24}$	15.079	3.503	47.01	1	0.6701
O_{3_8}	6.826	2.954	17.94	3	0.7794
PM_{10_24}	3.269	4.439	12.91	4	0.9291
$PM_{2.5\_24}$	5.026	3.154	14.11	5	1.1411
SO_{2_24}	1.834	4.694	3.16	2	0.2316
NO_{2_24}	11.926	4.775	20.89	2	0.6089
$CO_{\_24}$	0.156	5.336	0.30	1	0.203
O_{3_8}	11.825	5.931	25.72	3	0.8572
PM_{10_24}	19.685	4.078	29.45	4	1.0945
$PM_{2.5\_24}$	12.065	4.628	20.48	5	1.2048

TABLE 17 Subjective and objective weights of the pollutants.beijing

Pollutant	Indicator item				
	Indicator variability	Indicator conflict	Objective weight (%)	Subjective weight	Weight
SO_{2_24}	1.817	4.749	3.94	2	0.4394
NO_{2_24}	13.721	4.477	28.05	2	0.6805
$CO_{\_24}$	0.149	3.424	0.23	1	0.2023
O_{3_8}	11.461	6.522	34.14	3	0.9414
PM_{10_24}	15.91	3.437	24.98	4	1.0498
$PM_{2.5\_24}$	5.29	3.583	8.66	5	1.0866
SO_{2_24}	1.035	3.476	1.34	2	0.4134
NO_{2_24}	14.387	3.851	20.58	2	0.6058
$CO_{\_24}$	0.146	3.082	0.17	1	0.2017
O_{3_8}	11.098	6.171	25.44	3	0.8544
PM_{10_24}	15.413	3.027	17.33	4	0.9733
$PM_{2.5\_24}$	23.961	3.95	35.15	5	1.3515
SO_{2_24}	0.565	5.204	2.27	2	0.4227
NO_{2_24}	10.23	3.515	27.74	2	0.6774
$CO_{\_24}$	0.098	3.647	0.27	1	0.2027
O_{3_8}	5.291	6.808	27.78	3	0.8778
PM_{10_24}	12.656	3.336	32.56	4	1.1256
$PM_{2.5\_24}$	3.651	3.329	9.37	5	1.0937
SO_{2_24}	2.645	4.194	7.03	2	0.4703
NO_{2_24}	8.704	3.614	19.94	2	0.5994
$CO_{\_24}$	0.096	3.525	0.22	1	0.2022
O_{3_8}	14.689	5.058	47.11	3	1.0711
PM_{10_24}	9.653	3.142	19.23	4	0.9923
$PM_{2.5\_24}$	3.472	2.938	6.47	5	1.0647
SO_{2_24}	0.97	5.749	2.06	2	0.4206
NO_{2_24}	12.105	4.528	20.26	2	0.6026
$CO_{\_24}$	0.114	4.229	0.18	1	0.2018
O_{3_8}	8.016	6.858	20.32	3	0.8032
PM_{10_24}	13.578	4.058	20.36	4	1.0036
$PM_{2.5\_24}$	23.66	4.211	36.83	5	1.3683
SO_{2_24}	0.816	5.535	3.75	2	0.4375
NO_{2_24}	6.283	3.921	20.46	2	0.6046
$CO_{\_24}$	0.118	4.21	0.41	1	0.2041
O_{3_8}	4.934	7.215	29.57	3	0.8957
PM_{10_24}	6.778	3.895	21.93	4	1.0193
$PM_{2.5\_24}$	6.653	4.322	23.88	5	1.2388

TABLE 18 Subjective and objective weights of the pollutants.tianjin

Pollutant	Indicator item				
	Indicator variability	Indicator conflict	Objective weight (%)	Subjective weight	Weight
SO_{2_24}	3.514	3.983	8.81	2	0.4881
NO_{2_24}	8.205	3.637	18.78	2	0.5878
$CO_{\_24}$	0.227	3.529	0.50	1	0.205
O_{3_8}	6.49	7.159	29.24	3	0.8924
PM_{10_24}	10.755	3.635	24.60	4	1.046
$PM_{2.5\_24}$	8.5	3.377	18.07	5	1.1807
SO_{2_24}	3.723	4.747	4.79	2	0.4479
NO_{2_24}	12.772	4.37	15.13	2	0.5513
$CO_{\_24}$	0.523	4.241	0.60	1	0.206
O_{3_8}	16.067	5.435	23.67	3	0.8367
PM_{10_24}	39.322	4.331	46.16	4	1.2616
$PM_{2.5\_24}$	9.25	3.851	9.65	5	1.0965
SO_{2_24}	1.983	4.881	2.56	2	0.4256
NO_{2_24}	2.809	4.754	3.53	2	0.4353
$CO_{\_24}$	0.086	5.276	0.12	1	0.2012
O_{3_8}	5.114	4.755	6.43	3	0.6643
PM_{10_24}	54.359	5.187	74.55	4	1.5455
$PM_{2.5\_24}$	9.805	4.938	12.80	5	1.128
SO_{2_24}	1.155	5.006	4.58	2	0.4458
NO_{2_24}	6.69	5.681	30.12	2	0.7012
$CO_{\_24}$	0.15	3.692	0.44	1	0.2044
O_{3_8}	7.169	5.49	31.19	3	0.9119
PM_{10_24}	6.982	3.816	21.12	4	1.0112
$PM_{2.5\_24}$	4.363	3.629	12.55	5	1.1255
SO_{2_24}	3.232	3.712	8.90	2	0.489
NO_{2_24}	5.579	4.555	18.85	2	0.5885
$CO_{\_24}$	0.126	3.647	0.34	1	0.2034
O_{3_8}	8.549	3.675	23.31	3	0.8331
PM_{10_24}	13.842	2.615	26.84	4	1.0684
$PM_{2.5\_24}$	10.002	2.933	21.76	5	1.2176
SO_{2_24}	3.866	4.411	5.72	2	0.4572
NO_{2_24}	13.185	3.198	14.15	2	0.5415
$CO_{\_24}$	0.235	3.566	0.28	1	0.2028
O_{3_8}	5.698	6.026	11.52	3	0.7152
PM_{10_24}	42.799	2.861	41.07	4	1.2107
$PM_{2.5\_24}$	29.539	2.752	27.27	5	1.2727

TABLE 19 Corrected D-S fusion results.xi'an.

Pollutant					
SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
4.15628E-13	0.00895,313	0.002,902,783	0.011,033,686	0.015,980,774	0.016,979,121
2.78809E-13	0.008,647,582	0.002,865,373	0.010,689,599	0.01,545,038	0.017,715,433
4.18398E-13	0.009,892,757	0.004,150,316	0.013,571,875	0.040,238,703	0.027,800,006
1.90824E-13	0.007,681,425	0.002,949,608	0.014,901,633	0.018,394,637	0.017,312,711
2.87E-13	5.68E-03	9.49E-03	1.07E-02	1.02E-02	1.48E-02
2.71132E-13	0.008,670,017	0.002,898,068	0.012,197,168	0.015,434,065	0.017,061,811

TABLE 20 Corrected D-S fusion results.beijing

Pollutant					
SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
3.28354E-10	5.13756E-10	1.72723E-10	7.27037E-10	8.12909E-10	1.15087E-09
5.05213E-11	9.2492E-11	2.53405E-11	1.04744E-10	1.51516E-10	1.57875E-10
3.47249E-10	5.15963E-10	1.74465E-10	7.65038E-10	8.68307E-10	1.05426E-09
3.35082E-10	5.7305E-10	1.72864E-10	7.47497E-10	9.41563E-10	9.29306E-10
3.73908E-10	5.11179E-10	1.73433E-10	9.08838E-10	8.44144E-10	9.09717E-10
3.34311E-10	5.12807E-10	1.73045E-10	6.86848E-10	8.46532E-10	1.12507E-09

TABLE 21 Corrected D-S fusion results.tianjin

Pollutant					
SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
9.20799E-08	1.22106E-07	4.26766E-08	1.85607E-07	2.16034E-07	2.425E-07
8.44459E-08	1.13829E-07	4.26355E-08	1.71846E-07	2.40466E-07	2.27418E-07
8.08222E-08	9.16249E-08	4.2368E-08	1.39795E-07	3.22489E-07	2.34472E-07
8.44062E-08	1.46184E-07	4.27279E-08	1.90209E-07	2.10291E-07	2.3413E-07
9.2581E-08	1.22922E-07	4.25303E-08	1.73706E-07	2.19349E-07	2.51221E-07
8.66053E-08	1.13131E-07	4.25107E-08	1.49854E-07	2.34329E-07	2.62356E-07

dimensionless indices, facilitating robust comparisons among different pollutants across various regions (Tables 7–9). The daily pollutant concentrations, shown over the course of the year, varied notably between Xi'an, Beijing, and Tianjin (Figure 6).

3.3 Correlation analysis

Within a particular space, air pollution levels are influenced both by the concentration and toxicity of the pollutants and the interactions between different pollutants and regional environments. To test this hypothesis, we used SPSS matrix correlation analysis, which is widely recognized and utilized, simple to use, intuitive to interpret, and broadly applicable. Notably, this non-parametric method can robustly handle ordinal data, which is crucial in our research context.

We assessed the correlations using the correlation coefficient (Eq. (30)). This equation ensures reliable and statistically valid results, enabling us to gain deeper insights into the relationships and patterns within the data.

$$\rho_{12} = \frac{\mathrm{cov}(m_1, m_1)}{\sigma_{m_1}\sigma_{m_2}} = \frac{E\left(\left(m_1 - \mu_{m_1}\right)\left(m_2, \mu_{m_2}\right)\right)}{\sigma_{m_1}\sigma_{m_2}} \tag{30}$$

TABLE 22 Evaluation of the nearest neighbors in KNN.xi'an.

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.056,609,488	0.040,558,004	0.066,487,982	0.054,551,824	1
NO_2 level 2	0.042,155,944	0.151,585,617	0.196,875,714	0.130,205,758	
NO_2 level 3	0.190,304,093	0.390,670,328	0.507,058,834	0.362,677,752	
CO level 1	0.118,242,005	0.096,660,921	0.084,676,799	0.099,859,908	2
CO level 2	0.04,377,392	0.101,083,173	0.124,252,279	0.089,703,124	
CO level 3	0.129,984,945	0.481,969,556	0.548,436,671	0.386,797,058	
O_3 level 1	0.041,178,291	0.24,512,773	0.458,592,325	0.248,299,449	1
O_3 level 2	0.435,117,685	0.75,777,624	0.886,323,046	0.693,072,324	
O_3 level 3	0.7,002,692	1.127,437,411	1.245,273,238	1.024,326,617	
PM_{10} level 1	0.109,520,385	0.077,700,514	0.100,971,768	0.096,064,222	1
PM_{10} level 2	0.09,162,904	0.316,484,973	0.457,855,132	0.288,656,382	
PM_{10} level 3	0.599,634,053	0.741,587,328	0.883,630,931	0.741,617,437	
$PM_{2.5}$ level 1	0.113,124,541	0.079,991,266	0.113,333,526	0.102,149,778	1
$PM_{2.5}$ level 2	0.181,487,194	0.265,597,113	0.352,406,719	0.266,497,009	
$PM_{2.5}$ level 3	0.440,322,166	0.528,792,178	0.617,578,473	0.528,897,606	
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.831,836,401	0.799,592,627	0.768,817,483	0.80,008,217	3
NO_2 level 2	0.733,070,969	0.712,425,697	0.687,232,475	0.710,909,714	
NO_2 level 3	0.584,922,821	0.57,526,776	0.554,805,356	0.571,665,313	
CO level 1	0.261,714,568	0.238,188,844	0.217,802,665	0.239,235,359	2
CO level 2	0.187,246,482	0.190,368,447	0.185,305,735	0.187,640,222	
CO level 3	0.013,487,617	0.481,969,556	0.461,347,048	0.318,934,741	
O_3 level 1	0.917,444,775	0.780,772,073	0.685,078,468	0.794,431,772	2
O_3 level 2	0.523,505,381	0.696,175,417	0.722,294,137	0.647,324,978	
O_3 level 3	0.258,353,866	0.829,285,459	0.898,629,057	0.662,089,461	
PM_{10} level 1	1.494,750,669	1.445,338,396	1.397,798,937	1.445,962,667	3
PM_{10} level 2	1.293,601,244	1.219,234,497	1.147,764,378	1.220,200,039	
PM_{10} level 3	1.090,254,313	1.04,900,284	1.025,972,973	1.055,076,709	
$PM_{2.5}$ level 1	1.714,876,667	1.659,177,964	1.605,548,347	1.65,986,766	3
$PM_{2.5}$ level 2	1.547,575,525	1.493,219,886	1.442,381,952	1.494,392,454	
$PM_{2.5}$ level 3	1.395,446,239	1.352,818,968	1.314,919,222	1.35,439,481	
SO_2 level 1	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	1
SO_2 level 2	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	

(Continued on following page)

TABLE 22 (*Continued*) Evaluation of the nearest neighbors in KNN.xi'an.

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
NO_2 level 1	0.956,353,885	0.924,018,745	0.892,941,158	0.92,443,793	3
NO_2 level 2	0.857,588,453	0.835,114,723	0.808,675,472	0.833,792,883	
NO_2 level 3	0.709,440,305	0.68,673,658	0.654,206,705	0.683,461,196	
CO level 1	0.39,020,888	0.366,228,386	0.344,162,978	0.366,866,748	2
CO level 2	0.315,740,795	0.307,395,511	0.29,361,696	0.305,584,422	
CO level 3	0.14,198,193	0.481,969,556	0.411,128,385	0.345,026,624	
O_3 level 1	1.205,672,364	1.064,990,267	0.952,150,227	1.074,270,953	3
O_3 level 2	0.81,173,297	0.903,940,112	0.895,405,672	0.870,359,585	
O_3 level 3	0.546,581,455	0.933,638,585	0.971,320,177	0.817,180,072	
PM_{10} level 1	3.973,582,944	3.923,617,857	3.874,313,879	3.923,838,226	3
PM_{10} level 2	3.772,433,518	3.680,457,518	3.588,590,554	3.680,493,864	
PM_{10} level 3	3.500,093,981	3.415,229,814	3.33,427,542	3.416,533,072	
$PM_{2.5}$ level 1	2.72,333,398	2.667,269,329	2.612,460,109	2.667,687,806	3
$PM_{2.5}$ level 2	2.552,335,089	2.494,271,858	2.438,116,946	2.494,907,964	
$PM_{2.5}$ level 3	2.384,005,207	2.33,207,887	2.282,487,093	2.332,857,056	
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.735,220,646	0.703,070,053	0.6,726,076	0.703,632,767	3
NO_2 level 2	0.636,455,214	0.617,733,032	0.593,880,232	0.616,022,826	
NO_2 level 3	0.488,307,066	0.493,098,274	0.485,721,858	0.489,042,399	
CO level 1	0.2,701,381	0.246,568,072	0.22,601,261	0.247,572,927	2
CO level 2	0.195,670,015	0.197,604,316	0.191,643,799	0.19,497,271	
CO level 3	0.021,911,149	0.481,969,556	0.457,118,569	0.320,333,092	
O_3 level 1	1.338,648,154	1.19,676,297	1.079,042,028	1.204,817,717	3
O_3 level 2	0.94,470,876	1.013,160,666	0.993,458,963	0.98,377,613	
O_3 level 3	0.679,557,245	1.006,262,957	1.030,623,299	0.905,481,167	
PM_{10} level 1	1.789,176,344	1.739,615,972	1.691,594,142	1.740,128,819	3
PM_{10} level 2	1.588,026,918	1.508,553,784	1.430,853,337	1.50,914,468	
PM_{10} level 3	1.363,601,525	1.308,410,507	1.266,857,561	1.312,956,531	
$PM_{2.5}$ level 1	1.674,604,385	1.618,929,761	1.565,379,105	1.61,963,775	3
$PM_{2.5}$ level 2	1.507,553,451	1.453,455,946	1.402,995,156	1.454,668,184	
$PM_{2.5}$ level 3	1.35,657,697	1.314,629,677	1.277,593,731	1.316,266,793	
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.027,240,659	0.068,581,108	0.11,401,783	0.069,946,532	1
NO_2 level 2	0.126,006,091	0.206,396,724	0.252,789,344	0.195,064,053	
NO_2 level 3	0.274,154,239	0.453,178,439	0.568,909,964	0.432,080,881	

(Continued on following page)

TABLE 22 (*Continued*) Evaluation of the nearest neighbors in KNN.xi'an.

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
CO level 1	0.089,800,641	0.151,648,172	0.186,679,807	0.14,270,954	1
CO level 2	0.263,559,506	0.481,969,556	0.647,814,728	0.46,444,793	
CO level 3	0.015,332,556	0.04,720,816	0.081,772,845	0.04,810,452	
O_3 level 1	0.140,780,234	0.32,923,158	0.537,396,648	0.335,802,821	1
O_3 level 2	0.534,719,628	0.83,001,925	0.958,324,115	0.774,354,331	
O_3 level 3	0.799,871,143	1.198,804,543	1.316,557,244	1.105,077,643	
PM_{10} level 1	0.040,047,475	0.103,388,587	0.172,886,878	0.10,544,098	1
PM_{10} level 2	0.2,411,969	0.424,297,408	0.565,316,163	0.410,270,157	
PM_{10} level 3	0.706,822,133	0.848,571,614	0.99,046,006	0.848,617,936	
$PM_{2.5}$ level 1	0.04,187,612	0.11,367,406	0.192,183,295	0.115,911,158	1
$PM_{2.5}$ level 2	0.276,877,138	0.365,921,565	0.455,171,042	0.365,989,915	
$PM_{2.5}$ level 3	0.544,524,755	0.633,938,628	0.723,390,353	0.633,951,245	
SO_2 level 1	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	1
SO_2 level 2	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.834,079,878	0.801,834,207	0.771,052,741	0.802,322,275	2
NO_2 level 2	0.587,166,298	0.577,228,944	0.556,506,914	0.573,634,052	
NO_2 level 3	0.735,314,446	0.714,630,635	0.689,410,854	0.713,118,645	
CO level 1	0.190,516,019	0.193,165,568	0.187,746,342	0.190,475,976	1
CO level 2	0.264,984,104	0.241,440,818	0.22,098,739	0.242,470,771	
CO level 3	0.016,757,154	0.481,969,556	0.459,692,083	0.319,472,931	
O_3 level 1	0.409,110,763	0.873,620,695	0.926,232,573	0.736,321,344	1
O_3 level 2	0.674,262,278	0.798,637,341	0.804,648,908	0.759,182,842	
O_3 level 3	1.068,201,672	0.92,912,379	0.822,991,244	0.940,105,569	
PM_{10} level 1	1.493,119,192	1.443,707,908	1.396,171,694	1.444,332,931	2
PM_{10} level 2	1.08,877,026	1.047,616,982	1.024,716,105	1.053,701,116	
PM_{10} level 3	1.291,969,767	1.217,638,136	1.146,211,634	1.218,606,512	
$PM_{2.5}$ level 1	1.649,514,468	1.59,385,546	1.540,356,102	1.594,575,343	2
$PM_{2.5}$ level 2	1.332,403,273	1.290,901,487	1.254,431,847	1.292,578,869	
$PM_{2.5}$ level 3	1.482,626,266	1.428,697,126	1.378,482,774	1.429,935,389	

Figure 7 presents the correlations of the concentrations in each region. Figure 8 presents the correlations among different pollutant concentrations (Figure 8). There were complex interconnections and mutual influences among the diverse pollutants in terms of their concentrations in varying intervals. The dynamics of the air pollutants varied among the regions (Figure 7). This finding offers insight into how regional differences contribute to macro-scale air-pollution trends. Figure 8 illustrates the daily interactions among the pollutants. All of the correlation values exceed 4, indicating that they are statistically significant.

3.4 Case analysis

To verify the rationality and effectiveness of the air pollution data fusion method based on fuzzy comprehensive evaluation and DS evidence theory, this study utilized the original dataset provided by the China National Environmental Monitoring Center (CNEMC) and the Qingyue Open Environment Data Center (Tables 4–6). Systematic analysis and quantification (Tables 7–9) were conducted to consolidate air pollution indices from different regions for the same time period.

TABLE 23 Evaluation of the nearest neighbors in KNN.beijing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,014	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	
NO_2 level 1	0.131,687,243	0.210,802,444	0.257,127,561	0.199,872,416	2
NO_2 level 2	0.032,921,811	0.073,615,407	0.118,701,276	0.075,079,498	
NO_2 level 3	0.279,835,391	0.457,660,536	0.573,302,873	0.436,932,933	
CO level 1	0.024,822,695	0.055,505,233	0.0894,995	0.056,609,143	1
CO level 2	0.09,929,078	0.1,589,428	0.193,871,446	0.150,701,675	
CO level 3	0.273,049,645	0.481,969,556	0.655,337,954	0.470,119,052	
O_3 level 1	0.151,515,151	0.338,798,178	0.546,295,648	0.345,536,326	1
O_3 level 2	0.545,454,545	0.838,140,405	0.966,376,892	0.783,323,947	
O_3 level 3	0.810,606,061	1.206,735,193	1.324,458,238	1.113,933,164	
PM_{10} level 1	0.050,287,356	0.112,445,947	0.181,313,642	0.114,682,315	1
PM_{10} level 2	0.251,436,781	0.432,588,196	0.573,364,079	0.419,129,685	
PM_{10} level 3	0.714,717,621	0.85,636,282	0.998,175,522	0.856,418,654	
$PM_{2.5}$ level 1	0.056,666,667	0.126,710,519	0.204,314,572	0.129,230,586	1
$PM_{2.5}$ level 2	0.288,694,379	0.377,432,849	0.466,488,061	0.37,753,843	
$PM_{2.5}$ level 3	0.555,707,757	0.645,023,686	0.734,400,738	0.64,504,406	
SO_2 level 1	0.313,386,769	0.298,794,995	0.285,029,225	0.29,907,033	3
SO_2 level 2	0.253,519,332	0.223,585,614	0.220,175,824	0.232,426,923	
SO_2 level 3	0.096,367,312	0.000917,272	0.343,602,867	0.146,962,484	
NO_2 level 1	0.480,833,936	0.449,120,382	0.420,181,322	0.450,045,213	2
NO_2 level 2	0.1317	0.21,080,244	0.355,661	0.232,716,977	
NO_2 level 3	0.233	0.31,781,801	0.368,536	0.306,441,699	
CO level 1	0.147,899,935	0.125,555,459	0.110,084,606	0.127,846,667	2
CO level 2	0.07,343,185	0.110,550,934	0.126,372,269	0.103,451,684	
CO level 3	0.100,327,015	0.481,969,556	0.528,420,676	0.370,239,082	
O_3 level 1	0.575,521,889	0.450,264,983	0.407,527,949	0.477,771,607	2
O_3 level 2	0.181,582,495	0.552,939,714	0.63,636,368	0.456,961,963	
O_3 level 3	0.08,356,902	0.827,610,162	0.928,302,504	0.613,160,562	
PM_{10} level 1	0.762,621,255	0.714,106,716	0.669,642,365	0.715,456,778	2
PM_{10} level 2	0.56,147,183	0.525,020,578	0.503,730,411	0.530,074,273	
PM_{10} level 3	0.521,826,732	0.575,606,671	0.656,356,636	0.58,459,668	
$PM_{2.5}$ level 1	1.094,204,967	1.039,084,615	0.98,739,739	1.040,228,991	3
$PM_{2.5}$ level 2	0.933,154,078	0.885,500,181	0.844,682,812	0.887,779,024	
$PM_{2.5}$ level 3	0.811,733,934	0.787,641,638	0.773,234,224	0.790,869,932	
SO_2 level 1	0.014,966,859	0.033,466,914	0.053,963,778	0.034,132,517	1
SO_2 level 2	0.074,834,295	0.104,768,013	0.170,648,449	0.116,750,253	
SO_2 level 3	0.231,986,316	0.3,292,709	0.538,858,893	0.36,670,537	

(Continued on following page)

TABLE 23 (*Continued*) Evaluation of the nearest neighbors in KNN.beijing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
NO_2 level 1	0.03,292,181	0.073,615,406	0.118,701,276	0.075,079,497	1
NO_2 level 2	0.131,687,242	0.210,802,444	0.257,127,561	0.199,872,416	
NO_2 level 3	0.27,983,539	0.457,660,535	0.573,302,873	0.436,932,933	
CO level 1	0.024,822,695	0.055,505,233	0.0894,995	0.056,609,143	1
CO level 2	0.09,929,078	0.1,589,428	0.193,871,446	0.150,701,675	
CO level 3	0.273,049,645	0.481,969,556	0.655,337,954	0.470,119,052	
O_3 level 1	0.151,515,151	0.338,798,178	0.546,295,647	0.345,536,325	1
O_3 level 2	0.545,454,545	0.838,140,405	0.966,376,892	0.783,323,947	
O_3 level 3	0.81,060,606	1.206,735,192	1.324,458,237	1.113,933,163	
PM_{10} level 1	0.050,287,355	0.112,445,946	0.181,313,641	0.114,682,314	1
PM_{10} level 2	0.251,436,781	0.432,588,195	0.573,364,078	0.419,129,685	
PM_{10} level 3	0.71,471,762	0.856,362,819	0.998,175,521	0.856,418,654	
$PM_{2.5}$ level 1	0.056,666,666	0.126,710,518	0.204,314,572	0.129,230,585	1
$PM_{2.5}$ level 2	0.288,694,378	0.377,432,848	0.46,648,806	0.377,538,429	
$PM_{2.5}$ level 3	0.555,707,756	0.645,023,685	0.734,400,738	0.64,504,406	
SO_2 level 1	0.022,423,972	0.016,721,705	0.030,861,364	0.02,333,568	1
SO_2 level 2	0.037,443,464	0.067,377,183	0.142,988,629	0.082,603,092	
SO_2 level 3	0.194,595,485	0.291,880,069	0.509,772,943	0.332,082,832	
NO_2 level 1	0.018,196,089	0.036,065,115	0.081,275,272	0.045,178,825	1
NO_2 level 2	0.080,569,343	0.173,844,103	0.220,301,831	0.158,238,426	
NO_2 level 3	0.228,717,492	0.418,380,857	0.534,650,438	0.393,916,262	
CO level 1	0.007,479,424	0.040,738,103	0.075,682,951	0.041,300,159	1
CO level 2	0.081,947,509	0.145,803,071	0.180,889,043	0.136,213,207	
CO level 3	0.255,706,374	0.481,969,556	0.641,628,662	0.459,768,197	
O_3 level 1	0.060,631,355	0.260,697,106	0.473,367,897	0.264,898,786	1
O_3 level 2	0.454,570,749	0.771,407,102	0.899,971,971	0.708,649,941	
O_3 level 3	0.719,722,264	1.141,043,592	1.258,893,552	1.039,886,469	
PM_{10} level 1	0.034,127,017	0.052,820,212	0.12,054,281	0.069,163,346	1
PM_{10} level 2	0.167,022,408	0.367,190,242	0.509,374,706	0.347,862,452	
PM_{10} level 3	0.651,580,846	0.793,797,013	0.936,018,637	0.793,798,832	
$PM_{2.5}$ level 1	0.034,305,073	0.060,919,225	0.138,166,242	0.077,796,847	1
$PM_{2.5}$ level 2	0.221,735,932	0.311,054,413	0.400,483,852	0.311,091,399	
$PM_{2.5}$ level 3	0.489,963,498	0.579,470,091	0.668,992,816	0.579,475,468	
SO_2 level 1	0.319,344,384	0.304,745,278	0.290,954,568	0.305,014,743	3
SO_2 level 2	0.259,476,947	0.229,543,229	0.225,433,568	0.238,151,248	
SO_2 level 3	0.102,324,927	0.005,040,343	0.341,947,841	0.149,771,037	
NO_2 level 1	0.479,884,773	0.448,173,778	0.419,243,912	0.449,100,821	2
NO_2 level 2	0.1317	0.21,080,244	0.355,661	0.232,716,977	
NO_2 level 3	0.233	0.31,781,801	0.368,536	0.306,441,699	

(Continued on following page)

TABLE 23 (*Continued*) Evaluation of the nearest neighbors in KNN.beijing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
CO level 1	0.148,222,082	0.125,871,263	0.110,372,229	0.128,155,191	2
CO level 2	0.073,753,996	0.11,069,296	0.126,433,301	0.103,626,752	
CO level 3	0.100,004,869	0.481,969,556	0.528,208,239	0.370,060,888	
O_3 level 1	0.535,332,911	0.412,641,387	0.381,827,011	0.443,267,103	1
O_3 level 2	0.141,393,518	0.547,785,851	0.637,646,321	0.44,227,523	
O_3 level 3	0.123,757,998	0.836,641,503	0.939,933,043	0.633,444,181	
PM_{10} level 1	0.796,244,551	0.747,650,289	0.702,902,409	0.748,932,416	2
PM_{10} level 2	0.595,095,126	0.554,771,028	0.52,830,418	0.559,390,112	
PM_{10} level 3	0.539,352,214	0.585,796,387	0.660,208,371	0.595,118,991	
$PM_{2.5}$ level 1	1.068,399,516	1.013,318,543	0.96,176,705	1.014,495,037	3
$PM_{2.5}$ level 2	0.907,792,639	0.860,633,606	0.820,571,371	0.862,999,205	
$PM_{2.5}$ level 3	0.788,688,146	0.76,600,591	0.753,356,197	0.769,350,084	
SO_2 level 1	0.332,282,096	0.317,668,013	0.303,826,544	0.317,925,551	3
SO_2 level 2	0.27,241,466	0.242,480,941	0.236,965,522	0.250,620,374	
SO_2 level 3	0.115,262,639	0.017,978,055	0.338,686,952	0.157,309,215	
NO_2 level 1	0.483,041,049	0.451,321,586	0.42,236,131	0.452,241,315	3
NO_2 level 2	0.384,275,617	0.375,221,304	0.358,462,099	0.372,653,007	
NO_2 level 3	0.236,127,469	0.31,932,639	0.369,134,398	0.308,196,086	
CO level 1	0.149,642,357	0.127,263,955	0.111,642,546	0.129,516,286	2
CO level 2	0.075,174,272	0.111,328,075	0.126,711,793	0.104,404,713	
CO level 3	0.098,584,594	0.481,969,556	0.527,272,976	0.369,275,709	
O_3 level 1	0.613,522,829	0.486,217,991	0.433,858,241	0.511,199,687	2
O_3 level 2	0.219,583,435	0.560,426,398	0.637,483,248	0.472,497,694	
O_3 level 3	0.04,556,808	0.820,790,936	0.918,787,912	0.595,048,976	
PM_{10} level 1	0.818,019,546	0.769,377,368	0.724,460,049	0.770,618,987	2
PM_{10} level 2	0.616,870,121	0.57,426,597	0.544,734,604	0.578,623,565	
PM_{10} level 3	0.551,499,107	0.593,319,335	0.663,600,401	0.602,806,281	
$PM_{2.5}$ level 1	0.997,590,959	0.942,629,108	0.891,490,883	0.94,390,365	3
$PM_{2.5}$ level 2	0.838,341,633	0.792,709,598	0.754,986,597	0.795,345,943	
$PM_{2.5}$ level 3	0.726,405,853	0.708,075,284	0.700,799,685	0.711,760,274	

TABLE 24 Evaluation of the nearest neighbors in KNN.tianjing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
SO_2 level 1	0.014,966,767	0.033,466,832	0.053,963,701	0.034,132,433	1
SO_2 level 2	0.074,834,203	0.104,767,922	0.170,648,377	0.116,750,167	
SO_2 level 3	0.231,986,224	0.329,270,808	0.53,885,882	0.366,705,284	
NO_2 level 1	0.032,921,689	0.073,615,297	0.118,701,175	0.075,079,387	2
NO_2 level 2	0.131,687,121	0.210,802,349	0.257,127,468	0.199,872,312	

(Continued on following page)

TABLE 24 (*Continued*) Evaluation of the nearest neighbors in KNN.tianjing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
NO_2 level 3	0.279,835,269	0.457,660,439	0.573,302,778	0.436,932,829	
CO level 1	0.024,822,652	0.055,505,195	0.089,499,464	0.056,609,104	1
CO level 2	0.099,290,737	0.158,942,767	0.193,871,413	0.150,701,639	
CO level 3	0.273,049,603	0.481,969,556	0.65,533,792	0.470,119,026	
O_3 level 1	0.54,545,436	0.838,140,264	0.966,376,752	0.783,323,792	2
O_3 level 2	0.151,514,966	0.338,798,012	0.546,295,493	0.345,536,157	
O_3 level 3	0.810,605,875	1.206,735,055	1.324,458,101	1.11,393,301	
PM_{10} level 1	0.05,028,714	0.112,445,754	0.181,313,462	0.114,682,119	1
PM_{10} level 2	0.251,436,566	0.43,258,802	0.573,363,908	0.419,129,498	
PM_{10} level 3	0.714,717,454	0.856,362,655	0.998,175,359	0.856,418,489	
$PM_{2.5}$ level 1	0.056,666,424	0.126,710,302	0.204,314,371	0.129,230,365	1
$PM_{2.5}$ level 2	0.288,694,183	0.377,432,658	0.466,487,874	0.377,538,238	
$PM_{2.5}$ level 3	0.555,707,572	0.645,023,503	0.734,400,557	0.645,043,877	
SO_2 level 1	0.612,472,845	0.515,188,261	0.532,036,473	0.553,232,527	1
SO_2 level 2	0.769,624,866	0.739,691,148	0.717,448,218	0.742,254,744	
SO_2 level 3	0.829,492,302	0.814,662,939	0.800,118,714	0.814,757,985	
NO_2 level 1	1.105,366,352	1.072,949,738	1.041,605,919	1.073,307,336	2
NO_2 level 2	0.858,452,772	0.825,053,904	0.782,664,951	0.822,057,209	
NO_2 level 3	1.00660,092	0.982,543,911	0.955,049,965	0.981,398,265	
CO level 1	0.401,532,374	0.377,526,619	0.355,371,797	0.378,143,597	2
CO level 2	0.327,064,289	0.318,133,054	0.303,916,765	0.316,371,369	
CO level 3	0.153,305,423	0.481,969,556	0.408,348,675	0.347,874,552	
O_3 level 1	1.566,949,046	1.423,520,267	1.299,737,648	1.430,068,987	3
O_3 level 2	1.173,009,652	1.211,801,257	1.177,887,605	1.187,566,171	
O_3 level 3	0.907,858,137	1.156,254,011	1.16,124,607	1.075,119,406	
PM_{10} level 1	2.354,368,872	2.30,463,022	2.256,037,098	2.305,012,063	3
PM_{10} level 2	2.153,219,447	2.067,987,152	1.983,554,754	2.068,253,784	
PM_{10} level 3	1.906,008,697	1.836,221,647	1.775,108,958	1.839,113,101	
$PM_{2.5}$ level 1	2.217,512,352	2.161,588,579	2.107,228,935	2.162,109,955	3
$PM_{2.5}$ level 2	2.047,913,844	1.991,230,661	1.93,703,186	1.992,058,788	
$PM_{2.5}$ level 3	1.885,531,689	1.836,957,136	1.791,546,185	1.83,801,167	
SO_2 level 1	0.57,623,542	0.478,950,836	0.504,071,873	0.519,752,709	1
SO_2 level 2	0.73,338,744	0.703,453,722	0.681,619,786	0.706,153,649	
SO_2 level 3	0.793,254,877	0.778,431,914	0.76,390,786	0.77,853,155	
NO_2 level 1	0.883,326,864	0.851,042,068	0.820,130,621	0.851,499,851	3
NO_2 level 2	0.784,561,432	0.763,088,232	0.737,327,222	0.761,658,962	
NO_2 level 3	0.636,413,284	0.620,761,114	0.594,764,283	0.617,312,894	
CO level 1	0.324,389,632	0.315,593,107	0.301,477,285	0.313,820,008	1
CO level 2	0.150,630,767	0.481,969,556	0.408,978,678	0.347,193,001	

(Continued on following page)

TABLE 24 (*Continued*) Evaluation of the nearest neighbors in KNN.tianjing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
CO level 3	0.398,857,717	0.374,857,792	0.352,723,566	0.375,479,692	
O_3 level 1	1.246,437,568	1.105,356,024	0.990,880,736	1.114,224,776	3
O_3 level 2	0.852,498,174	0.936,770,299	0.924,537,901	0.904,602,125	
O_3 level 3	0.587,346,659	0.95,452,167	0.987,977,376	0.843,281,902	
PM_{10} level 1	3.174,600,904	3.124,718,222	3.075,671,033	3.12,499,672	3
PM_{10} level 2	2.973,451,479	2.883,858,764	2.794,560,926	2.883,957,056	
PM_{10} level 3	2.709,789,892	2.62,998,343	2.555,606,682	2.631,793,335	
$PM_{2.5}$ level 1	2.288,053,493	2.232,106,243	2.177,671,283	2.23,261,034	3
$PM_{2.5}$ level 2	2.118,219,613	2.061,302,393	2.006,769,848	2.062,097,285	
$PM_{2.5}$ level 3	1.954,821,561	1.905,668,884	1.859,533,516	1.906,674,653	
SO_2 level 1	0.014,966,775	0.033,466,839	0.053,963,708	0.03,413,244	1
SO_2 level 2	0.074,834,211	0.104,767,929	0.170,648,383	0.116,750,174	
SO_2 level 3	0.231,986,232	0.329,270,816	0.538,858,826	0.366,705,291	
NO_2 level 1	0.032,921,665	0.073,615,276	0.118,701,155	0.075,079,365	1
NO_2 level 2	0.131,687,097	0.21,080,233	0.257,127,449	0.199,872,292	
NO_2 level 3	0.279,835,245	0.45,766,042	0.57,330,276	0.436,932,808	
CO level 1	0.024,822,652	0.055,505,195	0.089,499,464	0.056,609,104	1
CO level 2	0.099,290,737	0.158,942,767	0.193,871,413	0.150,701,639	
CO level 3	0.273,049,603	0.481,969,556	0.65,533,792	0.470,119,026	
O_3 level 1	0.151,514,961	0.338,798,008	0.54,629,549	0.345,536,153	1
O_3 level 2	0.545,454,355	0.838,140,261	0.966,376,749	0.783,323,788	
O_3 level 3	0.81,060,587	1.206,735,052	1.324,458,097	1.113,933,007	
PM_{10} level 1	0.050,287,146	0.112,445,759	0.181,313,467	0.114,682,124	1
PM_{10} level 2	0.251,436,571	0.432,588,025	0.573,363,913	0.419,129,503	
PM_{10} level 3	0.714,717,458	0.85,636,266	0.998,175,363	0.856,418,494	
$PM_{2.5}$ level 1	0.056,666,433	0.126,710,309	0.204,314,377	0.129,230,373	1
$PM_{2.5}$ level 2	0.28,869,419	0.377,432,665	0.46,648,788	0.377,538,245	
$PM_{2.5}$ level 3	0.555,707,579	0.645,023,509	0.734,400,563	0.645,043,884	
SO_2 level 1	0.01,774,672	0.012,186,998	0.112,918,074	0.047,617,264	1
SO_2 level 2	0.077,614,157	0.064,410,331	0.056,297,884	0.066,107,457	
SO_2 level 3	0.1,394,053	0.236,689,884	0.468,997,091	0.281,697,425	
NO_2 level 1	0.089,999,811	0.06,589,191	0.070,134,899	0.075,342,207	1
NO_2 level 2	0.008,765,621	0.138,128,099	0.180,728,192	0.109,207,304	
NO_2 level 3	0.156,913,769	0.368,148,393	0.484,274,969	0.33,644,571	
CO level 1	0.017,707,586	0.025,822,296	0.059,031,247	0.034,187,043	1
CO level 2	0.056,760,499	0.128,508,664	0.163,482,011	0.116,250,392	
CO level 3	0.230,519,365	0.481,969,556	0.622,042,264	0.444,843,728	
O_3 level 1	0.022,191,318	0.199,202,146	0.413,156,024	0.211,516,496	1
O_3 level 2	0.371,748,076	0.715,242,837	0.843,441,924	0.643,477,612	

(Continued on following page)

TABLE 24 (*Continued*) Evaluation of the nearest neighbors in KNN.tianjing

Pollutant	Neighbor 1	Neighbor 2	Neighbor 3	Average	Level
O_3 level 3	0.636,899,591	1.084,351,258	1.202,017,625	0.974,422,825	
PM_{10} level 1	0.562,374,208	0.703,334,095	0.84,472,128	0.703,476,528	2
PM_{10} level 2	0.169,061,528	0.128,981,092	0.121,678,744	0.139,907,121	
PM_{10} level 3	0.032,087,897	0.284,288,275	0.422,269,749	0.246,215,307	
$PM_{2.5}$ level 1	0.170,938,446	0.237,717,889	0.315,981,141	0.241,545,825	2
$PM_{2.5}$ level 2	0.194,553,876	0.149,077,139	0.139,431,779	0.161,020,931	
$PM_{2.5}$ level 3	0.399,027,333	0.484,402,713	0.571,063,577	0.484,831,208	
SO_2 level 1	0.851,085,969	0.836,253,056	0.821,697,661	0.836,345,562	3
SO_2 level 2	0.791,218,533	0.761,284,814	0.738,817,138	0.763,773,495	
SO_2 level 3	0.634,066,512	0.536,781,928	0.549,160,487	0.573,336,309	
NO_2 level 1	1.098,390,041	1.065,976,733	1.034,643,654	1.066,336,809	3
NO_2 level 2	0.999,624,609	0.97,563,099	0.948,178,819	0.974,478,139	
NO_2 level 3	0.85,147,646	0.818,494,539	0.77,648,706	0.81,548,602	
CO level 1	0.400,284,714	0.376,281,669	0.354,136,415	0.376,900,933	2
CO level 2	0.325,816,629	0.316,947,957	0.302,778,315	0.315,180,967	
CO level 3	0.152,057,764	0.481,969,556	0.408,640,498	0.347,555,939	
O_3 level 1	1.347,022,528	1.205,070,424	1.087,081,933	1.213,058,295	2
O_3 level 2	0.687,931,619	1.011,247,384	1.034,816,307	0.91,133,177	
O_3 level 3	0.953,083,134	1.020,227,789	0.999,904,147	0.99,107,169	
PM_{10} level 1	2.29,300,266	2.243,279,017	2.1,947,336	2.243,671,759	3
PM_{10} level 2	2.091,853,234	2.007,090,084	1.923,193,434	2.007,378,918	
PM_{10} level 3	1.846,455,106	1.7,778,023	1.718,204,491	1.780,820,632	
$PM_{2.5}$ level 1	2.566,897,862	2.510,870,718	2.456,180,641	2.511,316,407	3
$PM_{2.5}$ level 2	2.396,269,009	2.338,567,532	2.282,922,575	2.339,253,039	
$PM_{2.5}$ level 3	2.229,488,128	2.17,842,686	2.129,909,458	2.179,274,815	

3.4.1 Calculation of BPA using the interval distance similarity

Calculation of interval similarity is critical in assessing the similarity between two intervals of data, particularly in the context of the variability of pollution concentrations. Equations (14) and (15) (Wu and Mendel, 2009), which calculate the similarity between fluctuating concentration values, generate BPA values that encapsulate the relationships among interval data points for atmospheric pollutants (Graziani et al., 2019). This approach ensures the accuracy of these similarity measures and establishes the foundation for data fusion, fostering more robust analysis and a comprehensive understanding of pollution trends.

BPA values corresponding to the atmospheric pollutant concentrations on selected days are illustrated in Tables 10–12. These values, derived from the interval similarity calculations (equations (14) and (15), facilitate objective assessment of the regional pollution levels for each pollutant, following Qin and Xiao (2019). These BPA values quantify the pollution level and establish the basis for the advanced stages of data fusion. These empirical insights (Tables 10–12) are paramount, enabling a systematic comparison pollutant levels and thereby highlighting the broader importance and impact of each pollutant.

3.4.2 Sequential fusion via DS evidence theory

Regional pollution levels were estimated via sequential fusion (Figure 4) using DS evidence theory (Tables 13–15). The uncorrected fusion results deviate significantly from the CNEMC AQI values, yielding an error rate of 50%.

3.4.3 Subjective and objective weights

Relying on uncorrected fusion data can lead to misleading conclusions. To improve estimation accuracy, the influence of toxicity and the interactions between pollutants must be considered. We therefore calculated the subjective and objective weights of the pollutants using the CRITIC weighting method (Tables 16–18).

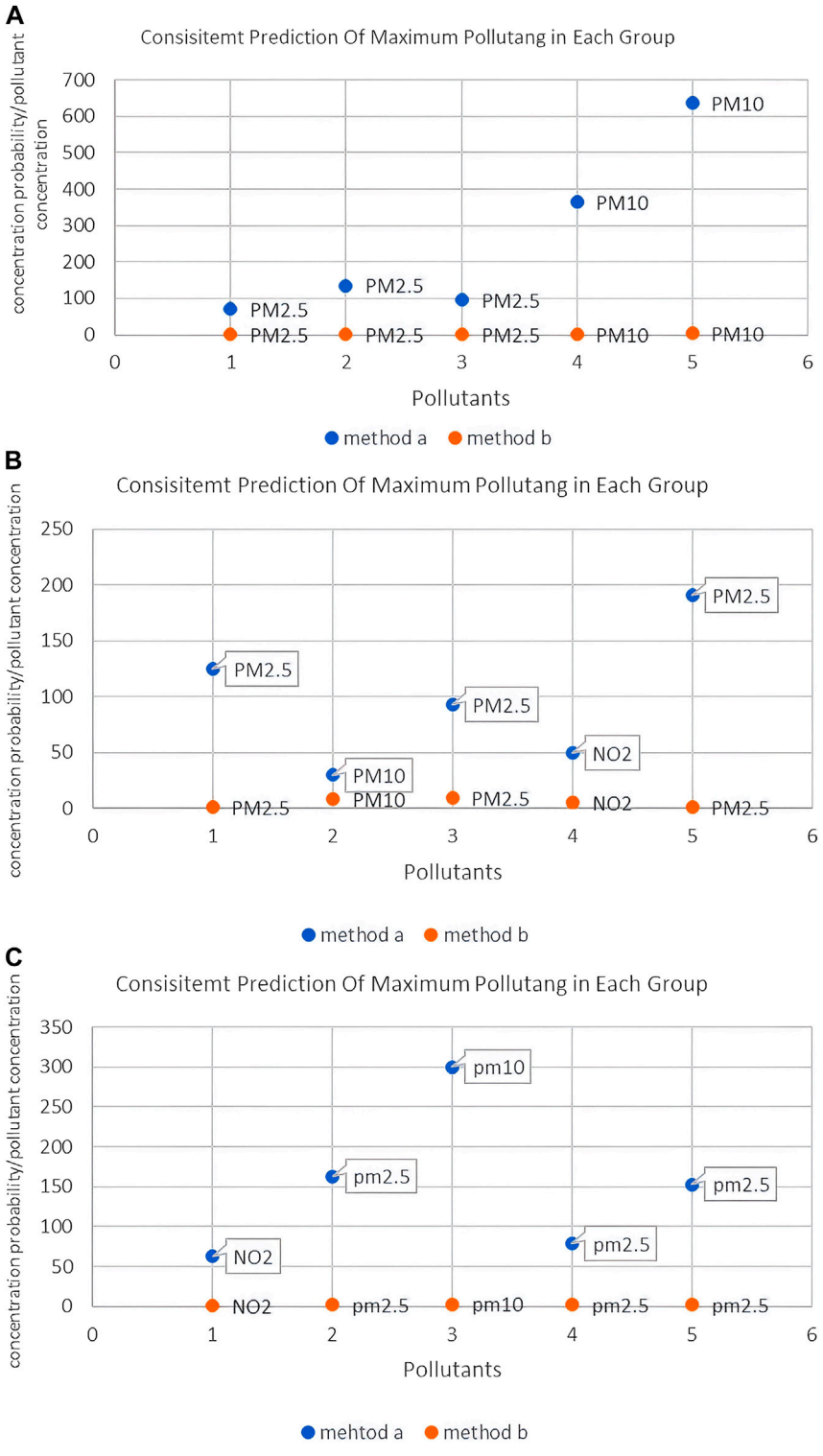


FIGURE 9
Comparison of predicted concentrations of major pollutants. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

3.4.4 DS fusion result correction using objective–subjective weights

The DS fusion results were corrected using objective–subjective weights (Tables 19–21). The high error rate in the uncorrected data (Table 7) underscores the need for careful interpretation and potential correction. Fusion-based correction achieved a remarkable and significant reduction in error, from the initial 50%–11%, suggesting a

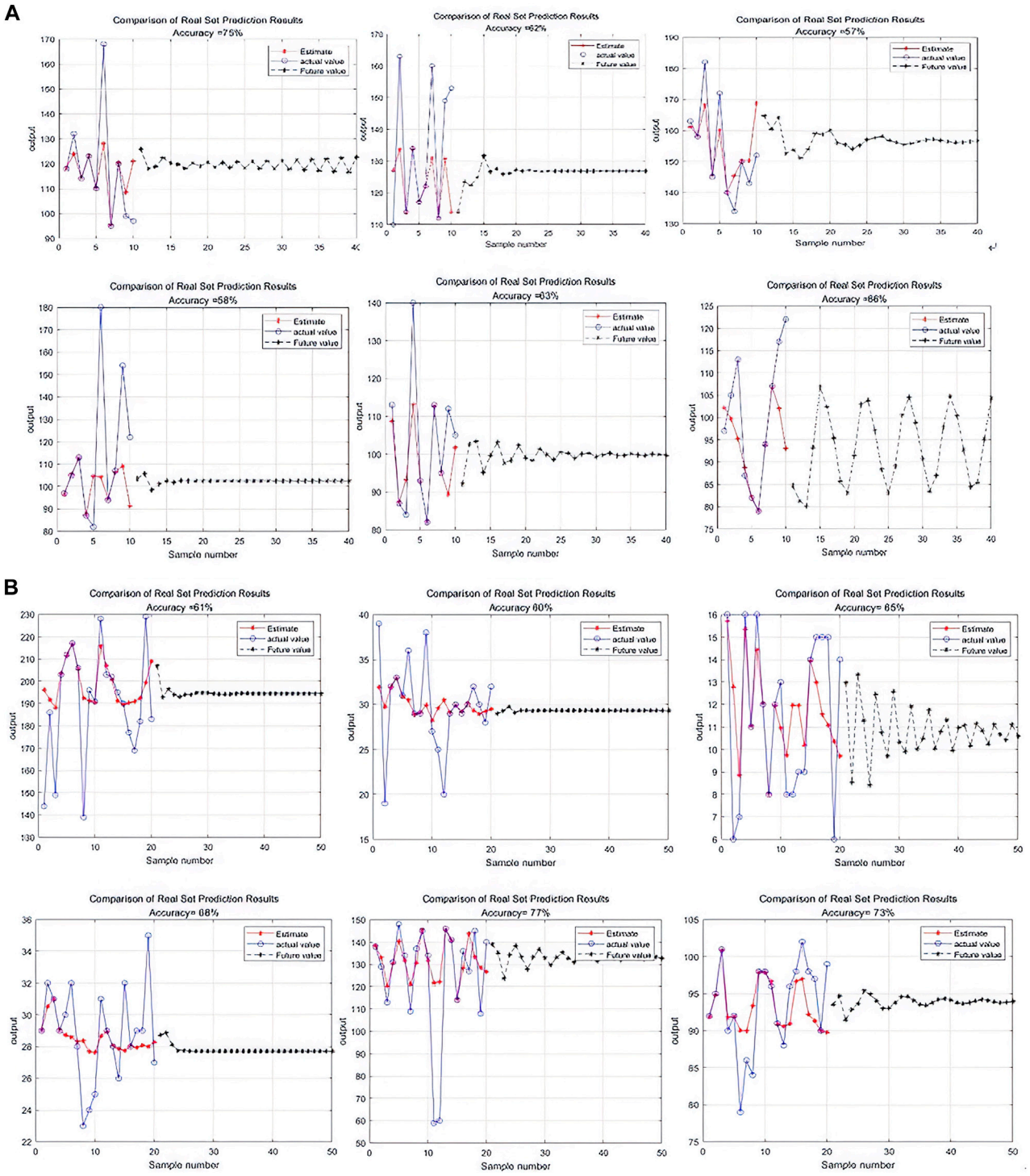


FIGURE 10
(Continued).

close correlation between the corrected results and the AQI values, reflecting the efficacy of the correction approach. This highlights the importance of addressing potential biases and uncertainty in the fusion process, generating more precise and reliable outcomes.

3.4.5 Concentration estimation via on the KNN approach and matter-element extension model

We used an extended KNN algorithm to estimate air pollutant levels, which were then correlated with air-quality levels. Tables 22–24 illustrates the variability in air quality (from good

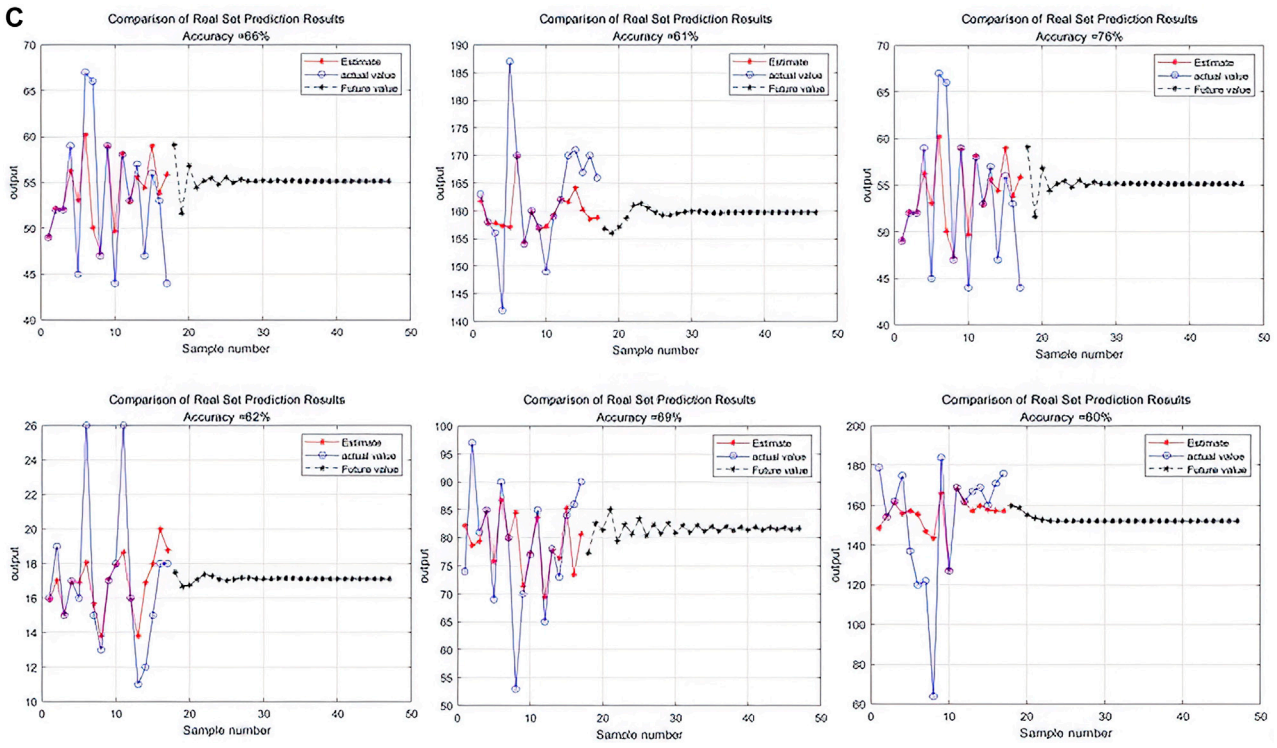


FIGURE 10
Comparison of support vector machine (SVM) training results. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

to severely polluted) during the study period. These results were consistent with the AQI values, facilitating the identification of major pollutants such as $PM_{2.5}$ and PM_{10} (Figure 9) and validating the accuracy and efficacy of the proposed method.

3.4.6 Further refinement of the DS fusion results

Given that air quality data are influenced by multiple factors, including meteorological conditions, vehicular traffic, and industrial activity, such data are inherently uncertain. Exclusive reliance on the raw DS fusion results might overlook pivotal environmental and socio-economic variables.

Refining DS fusion outcomes extends beyond merely enhancing accuracy, by more comprehensive and accurately portraying air quality and thereby equipping policymakers with evidence-based recommendations. For instance, a raw high pollution reading might trigger unwarranted alarm and panic, while a refined reading could suggest the necessity for additional monitoring sites or improved data fusion methodology. As air quality directly impacts public health, it is imperative to deliver precise data. Using raw values may cause the effects of pollutants to be under- or over-estimated, leading to misguided health advisories and policies. Here, via objective–subjective weighting, were able to integrate environmental and socio-economic aspects into the analysis, achieving more accurate and reliable outcomes. Such refinement is not limited to air quality assessment. As this method can handle uncertainty and complexity, it is applicable in other contexts, such as water and soil quality evaluations, highlighting the versatility and value of this novel approach.

The refinement of DS fusion outcomes were comprehensively validated using expanded datasets from urban environments in Xi'an, Beijing, and Tianjin, cities with distinct environmental and socioeconomic profiles. These findings highlight both the robustness of this refinement process and its applicability in diverse air quality scenarios. By integrating data from these cities, the method can provide substantially more refined and precise air quality assessments. This is critical for informed policy-making and effective public health management.

3.5 Comparison with SVM

To establish the effectiveness of the proposed method, we compared it to SVM. For our dataset, SVM achieved 75% prediction accuracy (Figure 10), whereas our method performed significantly better, achieving 86% accuracy (Figure 11). Further, unlike traditional SVM approaches, our method exhibited remarkable robustness and stability in handling uncertainty and noise. These findings validate the efficacy of our method and highlight its potential for evaluating trends in air pollution.

4 Discussion

In the current context of environmental science, precise estimation is essential for effective policy development. Our novel method aims to characterize air pollution dynamics with greater accuracy by integrating the evidence theory and the KNN algorithm.

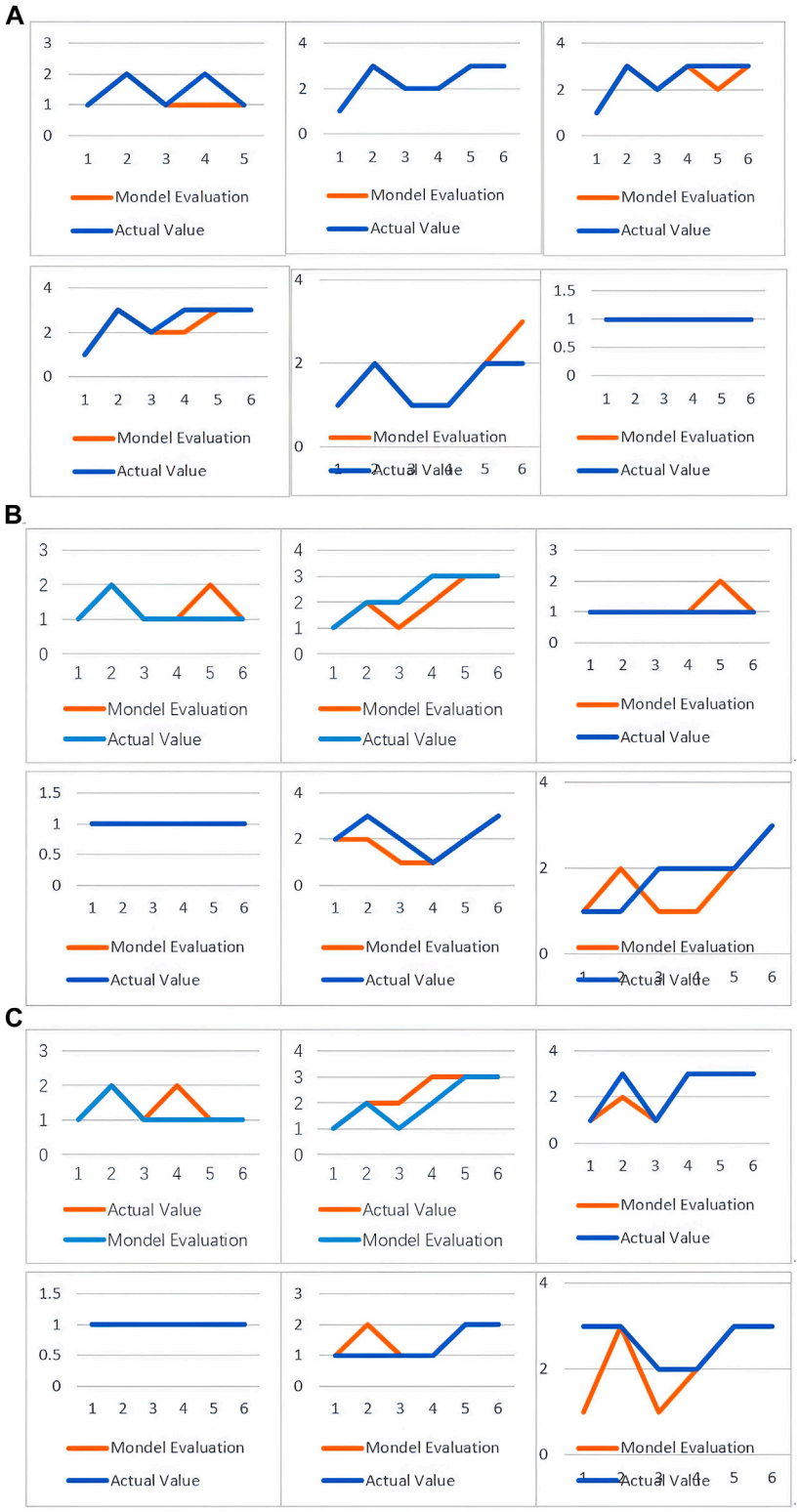


FIGURE 11
Comparison between values based on fuzzy evaluation of atmospheric pollution concentrations using evidence theory with KNN on the one hand and actual values on the other. **(A)** Xi'an; **(B)** Beijing; **(C)** Tianjin.

BPA, a cornerstone of the DS theory of evidence, effectively addresses the uncertainty inherent in atmospheric pollution data; we therefore used it for decision-making under uncertainty. Consistent with the findings of Zhong and Liu (2019), integrating BPA into our data fusion strategy reduced the uncertainty, enhancing the precision and reliability of the results.

Given the complexity of atmospheric pollution, a single monitoring point cannot capture the full extent of pollutant diffusion and distribution. Sequential fusion of atmospheric pollution concentrations, however, integrates multi-source monitoring data, estimates pollution levels more accurately and reliably, enhances data accuracy and reliability, and comprehensively describes the spatiotemporal variation in air pollution. By considering spatial correlation, this method more precisely assesses pollutant concentrations. This method achieves several key innovations, in terms of its strategy for handing small datasets, interdisciplinary approach, improved pollution atmospheric monitoring and prediction, and ability to consider multiple environments.

4.1 Innovative strategy for small samples

Limited data availability poses a primary challenge in environmental science, and especially in atmospheric pollution. Therefore, the effective analysis of small samples is important in improving the universality and applicability of models. Here, we aimed to achieve efficient and accurate air quality assessment with limited data availability. Consequently, we did not use the commonly applied neural network approach, primarily because it typically requires large datasets for effective training, and may not achieve optimal performance for small samples. By integrating evidence theory and the KNN approach, our method can handle limited data. Compared to other methods, this approach is more adaptable and precise, extracting meaningful insights from small datasets.

4.2 Interdisciplinary approach

By integrating evidence theory and the KNN algorithm, this model applies an interdisciplinary strategy. It represents a paradigm shift in atmospheric pollution data processing, with applicability in other environmental monitoring domains.

4.3 Enhanced monitoring and prediction

Compared to the traditional AQI approach, our model yields more comprehensive and precise assessments, by integrating multiple data sources and considering the characteristics of multiple pollutants. It thus comprehensively characterizes air pollution-related dynamics, especially for regions with limited data. This approach lays a robust foundation for effective policy development.

4.4 Data quality and methodological accuracy

This hybrid method, integrating interval similarity and subjective–objective weighting, accurately accounts for the inherent variability in pollutant concentrations and toxicological characteristics. By emphasizing data integrity and credibility, this approach will facilitate accurate and reliable assessments.

4.5 Expansion to multiple environments

Expanding our data to include Xi'an, Beijing, and Tianjin, which have varied urban settings, substantially strengthened the robustness and applicability of our findings, demonstrating the model's adaptability and effectiveness across different geographical landscapes. Such broad applicability reveals its potential for accurate and comprehensive air quality assessment for diverse urban contexts. This will enable it to become an essential tool for policymakers and environmental scientists.

4.6 Limitations

Although this method achieved significant results using a small sample data, this dataset was limited in its scope and diversity, focusing on specific urban environments and particular pollutants. This may restrict the model's applicability under a broader range of geographical and environmental conditions. The model does not currently adequately incorporate key external factors such as meteorological conditions, regional industrial activity, and traffic flow, which significantly impact air quality. While the model performs well with historical data, its ability to process real-time data and predict future trends remains to be validated. The current model's complexity may limit its direct application by non-professional users such as policymakers and environmental management agencies.

4.7 Future directions

Data from a wider range of geographical areas and more complex environmental and meteorological factors should be included to enhance the model's generalizability and accuracy. New and optimized algorithms should be used to enhance computational efficiency and accuracy. To validate its cross-field applicability, the method should be applied in environmental monitoring fields such as water quality assessment or forest health monitoring. Long-term research is required to better understand how pollutants change over time and under varying conditions. Finally, research findings should be translate into concrete policy recommendations and practical guidelines to promote environmental protection and sustainable development.

4.8 Conclusion

Our data-fusion evaluation method, which considers spatiotemporal variation, facilitates integration of multi-pollutant measurement data, provides a more granular view of air quality, and improves the evaluation of pollution levels and potential health impacts. It thus provides a tool for air quality evaluation during environmental assessments and in formulating public health policies.

Despite this significant progress, there remains broad scope for future development, particularly in terms of expanding the dataset, diversifying the model's applications, and optimizing the user interface. These improvements are anticipated to enhance the

model's role in air quality assessment and in formulating environmental management policies.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

All authors contributed to the study conception and design. CB analyzed the data and wrote the Introduction. GH reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (71874134) and Shaanxi Natural Science Basic Research Plan - Key Project (2019JZ-30).

Acknowledgments

We thank Qingyue Data (data.epmap.org) for providing the environmental data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ai, J., Niu, H., Chen, Z., Li, X., Zeng, Y., Zhuang, X., et al. (2022). D-S multi-information fusion GIL partial discharge joint recognition method based on neural network. *High. Volt. Eng.*, 481–488. doi:10.13336/j.1003-6520.hve.20211249

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* 46, 175–185. doi:10.1080/00031305.1992.10475879

Carslaw, D. C., and Rhys-Tyler, G. (2013). New insights from comprehensive on-road measurements of NO_x, NO_2 and NH_3 from vehicle emission remote sensing in London, UK. *Atmos. Environ.* 81, 339–347. doi:10.1016/j.atmosenv.2013.09.026

Chen, G., and Zhu, Y. (2014). An overview on the control of urban air pollution. *Sci. Technol. Rev.* 32, 15–22.

Chen, M., Qin, Q., Liu, F., Wang, Y., Wu, C., Yan, Y., et al. (2022). How long-term air pollution and its metal constituents affect type 2 diabetes mellitus prevalence? Results from Wuhan Chronic Disease Cohort. *Environ. Res.* 212, 113158. doi:10.1016/j.envres.2022.113158

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory.* 13, 21–27. doi:10.1109/TIT.1967.1053964

Cui, J., Sang, M., Bo, X., Wang, P., Xue, X. D., Guo, J., et al. (2022). Rapid assessment of optimization schemes for atmospheric emission sources based on integer programming algorithm. *Environ. Eng.* 40 (04), 202–208. doi:10.13205/j.hjgc.202204029

Dai, P., Yuming, W., and Feng, Z. (2018). An improved KNN algorithm based on analytic hierarchy process. *J. Liaoning Univ. Petrol. Chem. Technol.* 38, 87.

Dempster, A. P. (2008). "Upper and lower probabilities induced by a multivalued mapping," in *Classic works of the dempster-shafer theory of belief functions*. Editors R. Yager, and L. Liu (Berlin, Heidelberg: Springer), 57–72. doi:10.1007/978-3-540-44792-4_3

Dyer, R. F., and Forman, E. H. (1992). Group decision support with the analytic hierarchy process. *Decis. Support Syst.* 8, 99–124. doi:10.1016/0167-9236(92)90003-8

Eren, B., Aksangür, İ., and Erden, C. (2023). Predicting next hour fine particulate matter ($PM_{2.5}$) in the Istanbul metropolitan city using deep learning algorithms with time windowing strategy. *Urban Clim.* 48, 101418. doi:10.1016/j.uclim.2023.101418

Feng, X., Wang, B., Gao, G., Gao, S., Xie, C., and Shi, J. W. (2023). $Mn_yCo_{3-y}O_x$ bimetallic oxide prepared by ultrasonic technology for significantly improved catalytic performance in the reduction of NOx with NH_3. *Fuel* 352, 129159. doi:10.1016/j.fuel.2023.129159

Graziani, N. S., Carreras, H., and Wannaz, E. (2019). Atmospheric levels of BPA associated with particulate matter in an urban environment. *Heliyon* 5, e01419. doi:10.1016/j.heliyon.2019.e01419

He, K., Wang, T., Zhang, F., and Jin, X. (2022). Anomaly detection and early warning via a novel multiblock-based method with applications to thermal power plants. *Measurement* 193, 110979. doi:10.1016/j.measurement.2022.110979

Huihui, X., and Yanming, D. (2013). Improved the KNN algorithm based on related to the distance of attribute value. *Comput. Sci.* 40, 157–159.

Karavas, Z., Karayannis, V., and Moustakas, K. (2021). Comparative study of air quality indices in the European Union towards adopting a common air quality index. *Energy Environ.* 32 (6, 959–980. doi:10.1177/0958305X20921800

Kong, L., Wang, L., Li, F., Li, J., Wang, Y., Cai, Z., et al. (2023). Life cycle-oriented low-carbon product design based on the constraint satisfaction problem. *Energy Convers. Manag.* 286, 117069. doi:10.1016/j.enconman.2023.117069

Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, , et al. (2003). Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmos. Environ.* 37, 4539–4550. doi:10.1016/S1352-2310(03)00583-1

Li, G. D., Yamaguchi, D., and Nagai, M. (2007). A grey-based decision-making approach to the supplier selection problem. *Math. Comput. Modell.* 46, 573–581. doi:10.1016/j.mcm.2006.11.021

Li, J., Zhang, G., Ned, J. P., and Sui, L. (2023). How does digital finance affect green technology innovation in the polluting industry? Based on the serial two-mediator model of financing constraints and research and development (R&D) investments. *Environ. Sci. Pollut. Res. Int.* 30 (29), 74141–74152. doi:10.1007/s11356-023-27593-y

Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G., et al. (2023). Improving air quality assessment using physics-inspired deep graph learning. *npj Clim. Atmos. Sci.* 6, 152. doi:10.1038/s41612-023-00475-3

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., et al. (2016). Geospatial big data handling theory and methods: a review and research challenges. *ISPRS J. Photogramm.* 115, 119–133. doi:10.1016/j.isprsjprs.2015.10.012

Lü, C., Chen, S., Hua, X., and Xiang, W. (2021). Turbine through-flow fault diagnosis method and application based on the improved KNN algorithm. *Therm. Power Gener.* 50, 7. doi:10.19666/j.rlfd.202012289

Luo, J., Zhuo, W., Liu, S., and Xu, B. (2024). The optimization of carbon emission prediction in low carbon energy economy under big data. *IEEE Access* 12, 14690–14702. doi:10.1109/ACCESS.2024.3351468

Mo, H., Zhang, L., and Wang, H. (2020). A fuzzy comprehensive evaluation method for air quality assessment. *Environ. Monit. Assess.* 192, 243. doi:10.1007/s10661-020-8211-4

Murena, F. (2004). Measuring air quality over large urban areas: development and application of an air pollution index at the urban area of Naples. *Atmos. Environ.* 38, 6195–6202. doi:10.1016/j.atmosenv.2004.07.023

Pope, C. A., III, and Dockery, D. W. (2006). Health effects of fine particulate air pollution: lines that connect. *J. Air Waste Manag. Assoc.* 56, 709–742. doi:10.1080/10473289.2006.10464485

Pope, R. J., Coleman, P., and Patterson, K. G. (2022). The limitations of the air quality index: a critical review. *Environ. Sci. Policy.* 134, 55–65. doi:10.1007/978-981-10-5792-2_8

Priti, K., and Kumar, P. (2022). A critical evaluation of air quality index models (1960–2021). *Environ. Monit. Assess.* 194 (4), 324. doi:10.1007/s10661-022-09896-8

Qin, B., and Xiao, F. (2019). An improved method to determine basic probability assignment with interval number and its application in classification. *Int. J. Distrib. Sens. Netw.* 15, 155014771882052. doi:10.1177/1550147718820524

Ren, C., Niu, H. X., Chen, G. W., Si, Y. B., and She, Y. M. (2022). Heterogeneous data fusion of high-speed railway signal system based on DS evidence theory. *Transducer Microsyst. Technol.* 41, 54–57. doi:10.13873/J.1000-9787(2022)03-0054-04

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* 1, 83–98. doi:10.1504/IJSSCI.2008.017590

Seinfeld, J. H., and Pandis, S. N. (2016). *Atmospheric chemistry and physics: from air pollution to climate change*. Hoboken, NJ, USA: John Wiley & Sons.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ, USA: Princeton University Press.

Shang, K., Xu, L., Liu, X., Yin, Z., Liu, Z., Li, X., et al. (2023). Study of urban heat island effect in Hangzhou metropolitan area based on SW-TES algorithm and image dichotomous model. *SAGE Open* 13 (4). 21582440231208851 doi:10.1177/21582440231208851

Shang, M., and Luo, J. (2021). The Tapio decoupling principle and key strategies for changing factors of Chinese urban carbon footprint based on cloud computing. *Int. J. Environ. Res. Public Health* 18 (4), 2101. doi:10.3390/ijerph18042101

Shukla, K., Dadheech, N., Kumar, P., and Khare, M. (2021). Regression-based flexible models for photochemical air pollutants in the National Capital Territory of megacity Delhi. *Chemosphere* 272, 129611. doi:10.1016/j.chemosphere.2021.129611

Sowlat, M. H., Gharibi, H., Yunesian, M., Tayefeh Mahmoudi, M. T., and Lotfi, S. (2011). A novel, fuzzy-based air quality index (FAQI) for air quality assessment. *Atmos. Environ.* 45, 2050–2059. doi:10.1016/j.atmosenv.2011.01.060

Sun, Q., Zhang, T., Wang, X., and Chen, Z. (2022). Air quality evaluation method based on progressive DS evidence theory. *J. Beijing For. Univ.* 44, 119–128.

Sun, Q., Zhang, T., Wang, X., Lin, W., Fong, S., Chen, Z., et al. (2022). An ambient air quality evaluation model based on improved evidence theory. *Sci. Rep.* 12, 5753. doi:10.1038/s41598-022-09344-0

Tan, Q., Zhuang, H., Tie, Z. X., Ding, C. F., and Wu, Q. (2017). An air quality evaluation model for the industrial park based on the fuzzy matter-element analysis. *J. Saf. Environ.* 17, 2019–2023. doi:10.13637/j.issn.1009-6094.2017.05.076

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* 46, 234–240. doi:10.2307/143141

Utku, A., Can, Ü., Kamal, M., Das, N., Cifuentes-Faura, J., and Barut, A. (2023). A long short-term memory-based hybrid model optimized using a genetic algorithm for particulate matter 2.5 prediction. *Atmos. Pollut. Res.* 14, 101836. doi:10.1016/j.apr.2023.101836

Vaidya, O. S., and Kumar, S. (2006). Analytic hierarchy process: an overview of applications. *Eur. J. Oper. Res.* 169, 1–29. doi:10.1016/j.ejor.2004.04.028

Wang, J., Luo, Q., and Hu, Y. (2021). Locomotive bearing fault diagnosis method based on KNN-EMD algorithm. *Comput. Simul.* 38, 129–132.

Wang, W., and Klir, G. J. (2009). *Generalized measure theory*. New York, NY, USA: Springer Science and Business Media.

Wang, X., Wang, T., Xu, J., Shen, Z., Yang, Y., Chen, A., et al. (2022). Enhanced habitat loss of the Himalayan endemic flora driven by warming-forced upslope tree expansion. *Nat. Ecol. Evol.* 6 (7), 890–899. doi:10.1038/s41559-022-01774-3

Wang, Y., and Jiang, Y. (2017). Air quality evaluation based on improved CRITIC weighting and fuzzy optimization. *Stat. Decis.* 17, 83–87. doi:10.13546/j.cnki.tjyjc.2017.17.019

Wu, B., Gu, Q., Liu, Z., and Liu, J. (2023). Clustered institutional investors, shared ESG preferences and low-carbon innovation in family firm. *Technol. Forecast. Soc. Change* 194, 122676. doi:10.1016/j.techfore.2023.122676

Wu, D., and Mendel, J. M. (2009). A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets. *Inf. Sci.* 179, 1169–1192. doi:10.1016/j.ins.2008.12.010

Xiao, H., and Duan, Y. (2013). An improved study on the KNN algorithm based on the attribute value-related distance. *Comput. Sci.* 40, 4.

Xiong, L., Bai, X., Zhao, C., Li, Y., Tan, Q., Luo, G., et al. (2022). High-resolution data sets for global carbonate and silicate rock weathering carbon sinks and their change trends. *Earth's Future* 10 (8). e2022EF002746 doi:10.1029/2022EF002746

Yang, Y., Xiang, C. C., and Wei, D. J. (2010). Data cluster based on extension K nearest neighbor algorithm. *Comput. Eng. Appl.* 46, 156–159.

Zadeh, L. A. (1965). Fuzzy sets. *Inf. Control.* 8, 338–353. doi:10.1016/S0019-9958(65)90241-X

Zhang, B. (1997). *The fuzzy matter-element analysis*. Beijing, China: Petroleum Industry Press.

Zhang, S., Bai, X., Zhao, C., Tan, Q., Luo, G., Wang, J., et al. (2021). Global CO_2 consumption by silicate rock chemical weathering: its past and future. *Earth's Future* 9 (5), e1938E–e2020E. doi:10.1029/2020EF001938

Zhang, Y. L., and Cao, F. (2015). Fine particulate matter ($PM_{2.5}$) in China at a city level. *Sci. Rep.* 5, 14884. doi:10.1038/srep14884

Zhong, S., and Liu, X. (2019). *A new method to determine basic probability assignment based on interval number. Computing, Communications and IoT Applications*. Shenzhen, China: ComComAp, 316–320. doi:10.1109/ComComAp46287.2019.9018805

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 09 July 2024
DOI 10.3389/fenvs.2024.1399339

OPEN ACCESS

EDITED BY
Sayali Sandbhor,
Symbiosis International University, India

REVIEWED BY
Arabinda Maiti,
Vidyasagar University, India
Seyed Vahid Razavi-Termeh,
Sejong University, Republic of Korea

*CORRESPONDENCE
Mina Karimi,
mina.karimi@univie.ac.at

RECEIVED 11 March 2024
ACCEPTED 10 June 2024
PUBLISHED 09 July 2024

CITATION
Bashardoost A, Mesgari MS and Karimi M (2024), Enhancing spatial modeling and risk mapping of six air pollutants using synthetic data integration with convolutional neural networks.
Front. Environ. Sci. 12:1399339.
doi: 10.3389/fenvs.2024.1399339

Enhancing spatial modeling and risk mapping of six air pollutants using synthetic data integration with convolutional neural networks

Abed Bashardoost[1], Mohammad Saadi Mesgari[1] and Mina Karimi[1,2]*

[1]GIS Department, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran, [2]Department of Geography and Regional Research, University of Vienna, Vienna, Austria

Air pollution poses significant risks to human health and the environment, necessitating effective air quality management strategies. This study presents a novel approach to air quality management by integrating an autoencoder (AE) with a convolutional neural network (CNN) algorithm in Tehran city of Iran. One of the primary and vital problems in deep learning is model complexity, and the complexity of a model is affected by data distribution, data complexity, and information volume. AE provide a helpful way to denoise input data and make building deep learning models much more efficient. The proposed methodology enables spatial modeling and risk mapping of six air pollutants, namely, particulate matter 2.5 ($PM_{2.5}$), particulate matter 10 (PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and carbon monoxide (CO). For air pollution modelling, data from a spatial database containing the annual average of six pollutants from 2012 to 2022 was utilized. The model considered various parameters influencing air pollution: altitude, humidity, distance to industrial areas, NDVI (normalized difference vegetation index), population density, rainfall, distance to the street, temperature, traffic volume, wind direction, and wind speed. The risk map accuracy was assessed using the area under the receiver operating characteristic (ROC) curve for six pollutants. Among them, NO_2, PM_{10}, CO, $PM_{2.5}$, O_3, and SO_2 exhibited the highest accuracy with values of 0.964, 0.95, 0.896, 0.878, 0.877, and 0.811, respectively, in the risk map generated by the CNN-AE model. The findings demonstrated the CNN-AE model's impressive precision when generating the pollution risk map.

KEYWORDS

air pollutants, synthetic data, risk mapping, deep learning, spatial modeling, urban pollution

1 Introduction

Pollution occurs when toxic substances enter the environment and impact humans and other organisms. Pollutants refer to harmful substances (solids, liquids, or gases) present in higher-than-normal concentrations and harm the environment's quality (Manisalidis et al., 2020). Furthermore, increasing population growth in large cities has intensified air pollution, a significant environmental repercussion. This situation has been exacerbated

by various factors, including the rise in the utilization of heating devices, the presence of industrial centers, increased commercial activities, and the reliance on fossil fuels for transportation and traffic (Shogrkhodaei et al., 2021). Since cities contain a large population, they have the highest load of ambient air pollution. According to statistics announced by the World Health Organization (WHO), about 9 out of 10 people (approximately 91% of the world's population) are exposed to a high level of air pollution (Agarwal, 2021). Air pollution has many effects on human life, and poor air quality leads to the death of three million people annually (Sakti et al., 2023). The impact of air pollution includes an increase in cardiovascular and respiratory diseases, diabetes and blood pressure, dementia, the chance of miscarriage in pregnancy, early psychiatric and mental mortality, memory issues, the impairment of cognitive function, and a decrease in life expectancy. Severe air pollution can lead to an increase in criminal and immoral behavior, a reduction in the happiness of city dwellers, and a decrease in the potential of solar energy (Liu et al., 2020).

The uncontrolled expansion of urban areas and the rapid growth of industries have significantly diminished the quality of life and the environment in developing nations. Consequently, there is a pressing need to evaluate the geographical dispersion of air quality and its impact on human populations residing in metropolitan regions (Sengupta et al., 1996). The issues stemming from air pollution and their potential threat to human life underscore the significance of diligent air quality monitoring. Such monitoring is crucial for precise air quality regulation and effective urban management (Ma et al., 2019). It is essential to control and reduce ambient air pollutants, improve air quality, and maintain public health in urban areas. This is possible by developing appropriate strategies and policies and investigating and understanding the spatial changes of ambient air pollutants because being modifiable and reversible is one of the characteristics of ambient air pollution (Faridi et al., 2019). For analyzing air quality in a city, pollution maps, which show the average pollution level in a certain period, are considered a suitable tool (Szopińska et al., 2022).

It is necessary to understand the location to find the right solution for the population health problems caused by air pollution. The location of people in cities plays a vital role in exposure to air pollution (Zou et al., 2014). Therefore, the spatial analysis of air pollution can lead to the understanding of the location of pollutants in the city. Spatial analysis can help people solve complex location-based problems. The spatial analysis involves comprehending various aspects, including the distinctive features of a place, the interconnections between different locations, and the incidence of events within specific geographical areas (Farahani et al., 2022). It is possible to perform spatial analysis and solve spatial problems using a geographic information system (GIS) (Lü et al., 2019). In GIS, the first step in processing and analyzing any phenomenon is the spatial modeling of that phenomenon (Hogland and Anderson, 2017). GIS is a fundamental tool utilized in air pollution modeling. It enables the extraction and processing of spatial data necessary as input for air pollution models and the visualization of the models' outcomes (Makarovskikh and Herreinstein, 2022). GIS provides the results of urban air quality in the form of maps, which are very visual and can be easily interpreted even by non-specialists. He also analyzed these maps according to their complexity and user ability (Mavroulidou et al., 2004). Spatial analysis and overlay techniques available in GIS provide a powerful tool for pollution mapping (Briggs et al., 1997). GIS is essential for spatially monitoring air quality and creating spatial models to predict future air quality conditions ((Gulliver and Briggs, 2011; Somvanshi et al., 2019). Researchers use GIS techniques in various fields of air pollution investigation, such as analyzing air pollutants' spatial and temporal distribution (Kumar et al., 2016; Razavi-Termeh et al., 2021) and converting point data to surface data in studying the spatial distribution of air pollutants (Bell, 2006) have used.

So far, various methods such as land use regression (LUR), machine learning, and deep learning have been used to monitor and model air pollutants. Among the research conducted using the LUR method are: Mölter et al. (2010) estimated annual mean nitrogen dioxide (NO_2), and particulate matter 10 (PM_{10}) concentrations from 1996 to 2008 for Manchester using LUR models. Xu et al. (2019) investigated national particulate matter 2.5 (PM2.5) and NO_2 exposure models based on China's LUR, satellite measurements, and kriging method. Shi et al. (2020) studied a temporal LUR model for assessing ambient $PM_{2.5}$ in Pakistan. Xu et al. (2022) used the 3D LUR method to assess $PM_{2.5}$ exposure in central Taiwan. Ge et al. (2022) investigated the LUR method to determine exposure to ultrafine particles (UFP) in Shanghai, China. LUR models usually cannot be generalized to places other than the place developed for them, and optimizing the features for new models in specific study areas is a cumbersome process (Steininger et al., 2020). Another weakness of the LUR model is the need for experimental data (Dons et al., 2014). To address the limitations of the LUR method, machine learning algorithms have been developed to establish nonlinear and intricate relationships between observations and predictive variables. These algorithms offer several advantages, including rapid processing speed, higher efficiency compared to traditional models, and the absence of a requirement for statistical assumptions (Shogrkhodaei et al., 2021). One notable advantage of machine learning-based methods is their ability to operate without an in-depth understanding of atmospheric pollutants' physical or chemical properties (Bekkar et al., 2021). Machine learning methods generally work well and can identify data patterns quickly. Studies that have been used machine learning algorithms to investigate air pollution are described as follows: Hu et al. (2016) introduced a dense air pollution estimation model based on support vector regression (SVR) using a static and wireless sensor network. The results showed that air pollution estimations through this method have high spatial resolution and are more accurate than artificial neural network (ANN) model estimations. Delavar et al. (2019) investigated a new method to improve air pollution prediction based on machine learning approaches (SVR, geographically weighted regression, ANN, and auto-regressive nonlinear neural network with external input). According to their findings, the autoregressive nonlinear neural network that utilizes the proposed prediction model and external information is the most dependable algorithm for predicting air pollution. Castelli et al. (2020) investigated a machine-learning approach to predict air quality in California. The results indicated the possibility of predicting the concentration of carbon monoxide (CO), sulfur dioxide (SO_2), NO_2, ozone (O_3), and $PM_{2.5}$ as well as air quality

index (AQI) for the state of California with the SVR with radial basis function (RBF) kernel. Shogrkhodaei and colleagues (2021) conducted a study in Tehran, focusing on the spatial-temporal modeling of $PM_{2.5}$ risk mapping. Their analysis used three machine learning algorithms - random forest (RF), AdaBoost, and stochastic gradient descent (SGD) -. Abu El-Magd et al. (2022) employed a machine learning approach to develop a PM pollution susceptibility map using time series data of PM pollution records. The findings demonstrated that the generated prediction maps are reliable and could aid in enhancing air quality monitoring in the future.

Much research has been done on air pollution using deep learning algorithms in recent years. Deep learning algorithms have been preferred over machine learning models for greater flexibility and predictive accuracy (especially for big data) (Ghorbanzadeh et al., 2019). Deep learning, having advantages such as using more layers, more expansive data sets, and processing all layers simultaneously to obtain more accurate results, is suitable for modeling and forecasting air pollution (Bekkar et al., 2021). The following studies have been reviewed that investigated air pollution through the deep learning method. Bui et al. (2018) investigated a deep learning approach for forecasting air pollution using long short-term memory (LSTM) in South Korea. The proposed model showed significant results in predicting $PM_{2.5}$ in the long future based on historical meteorological data. Kalajdjieski et al. (2020) predicted air pollution with multi-modal data and deep neural networks (DNN). The results showed the substantial accuracy of this method, which was comparable to sequence models and conventional models that use air pollution data. Zaini et al. (2022) utilized a hybrid deep learning model to forecast hourly $PM_{2.5}$ concentration in an urban region of Malaysia. The outcomes revealed that the EEMD-LSTM model had the best accuracy compared to other deep learning models, and the combined prediction model outperformed the individual models. In the deep learning method, over fitting may happen with random noise in the data. Also, one of the primary and vital problems in deep learning is model complexity, and the complexity of a model is affected by data distribution, data complexity, and information volume (Hu et al., 2021). Autoencoders (AE) provide a helpful way to denoising input data and make building deep learning models much more efficient. Among the uses of AE is the ability to identify anomalies, deal with unsupervised learning problems, and remove complexity in data sets (Bank et al., 2020). Combining deep learning with AE in studies such as spatiotemporal modeling and prediction in cellular networks (Wang et al., 2017), traffic flow prediction with big data (Lv et al., 2014), and pollution map recovery with mobile sensing networks (Ma et al., 2019) had acceptable accuracy.

The innovation of this research lies in integrating AE with convolutional neural networks (CNN) for improved spatial modeling and risk mapping of air pollutants. This approach enhances predictive accuracy and efficiency by using AE to denoise and reduce data complexity while CNN captures complex spatial patterns. The combined CNN-AE model outperforms traditional methods like LUR and essential machine learning by automating feature extraction and handling large, complex datasets. The methodology generates high-resolution risk maps, aiding policymakers and public health officials in identifying pollution hotspots and implementing targeted interventions. This study significantly advances air pollution modeling and management by addressing the limitations of traditional models and leveraging advanced deep-learning techniques.

2 Materials and methods

2.1 Methodology

This research was generally conducted in six general steps (Figure 1). In the first step, six air pollutants, including $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3, and CO, were monitored over 10 years. During this phase, data from monitoring stations were collected to capture the concentrations of these pollutants. A comprehensive study area map was also created, incorporating 11 spatial criteria known to influence air pollutant levels. These spatial criteria included altitude, humidity, distance to industrial areas, NDVI (normalized difference vegetation index), population density, rainfall, distance to the street, temperature, traffic volume, wind direction, and wind speed. In the next step, the researchers aimed to determine the presence of multicollinearity among the spatial criteria. The multicollinearity test assessed the degree of correlation between the various parameters. This analysis was crucial in identifying redundant or highly correlated variables, allowing for eliminating or consolidating such factors to avoid multicollinearity issues in subsequent modeling steps. To understand the importance of the spatial criteria with air pollutant concentrations, the Geodetector method was employed. This method assessed the contributions and significance of each spatial criterion to the overall air pollution levels. It helped prioritize influential factors and provided insights into the relative importance of various spatial parameters. The researchers combined the CNN algorithm and the AE technique in the modeling phase. By integrating these two methods, the researchers could leverage the strengths of both approaches. The encoded data was then fed into the CNN, which effectively learned the spatial relationships and patterns associated with the concentrations of the six air pollutants. This fusion approach enhanced the accuracy and precision of the modeling process. The trained CNN-AE model generated risk maps for the six air pollutants. To generate risk maps, the predicted values obtained for each pixel in the study area were assigned to the center of each pixel. Then, using raster to point analysis, the risk map was created. In the next step, the natural breaks classification method was utilized to categorize the risk classes for each pollutant. These risk maps provided spatial representations of the pollutant concentrations across the study area. The results and risk maps obtained from the modeling process were evaluated and compared in the final step. Evaluation metrics such as mean absolute error (MAE), coefficient of determination (R^2), root mean square error (RMSE), and the area under the curve (AUC) of the receiver operating characteristic (ROC) were employed to assess the accuracy and performance of the models.

2.2 Study area

Tehran city (the capital of Iran) is located at 35° 36′to 35° 44′north latitude and 51° 17′to 51° 33′east longitude and an altitude

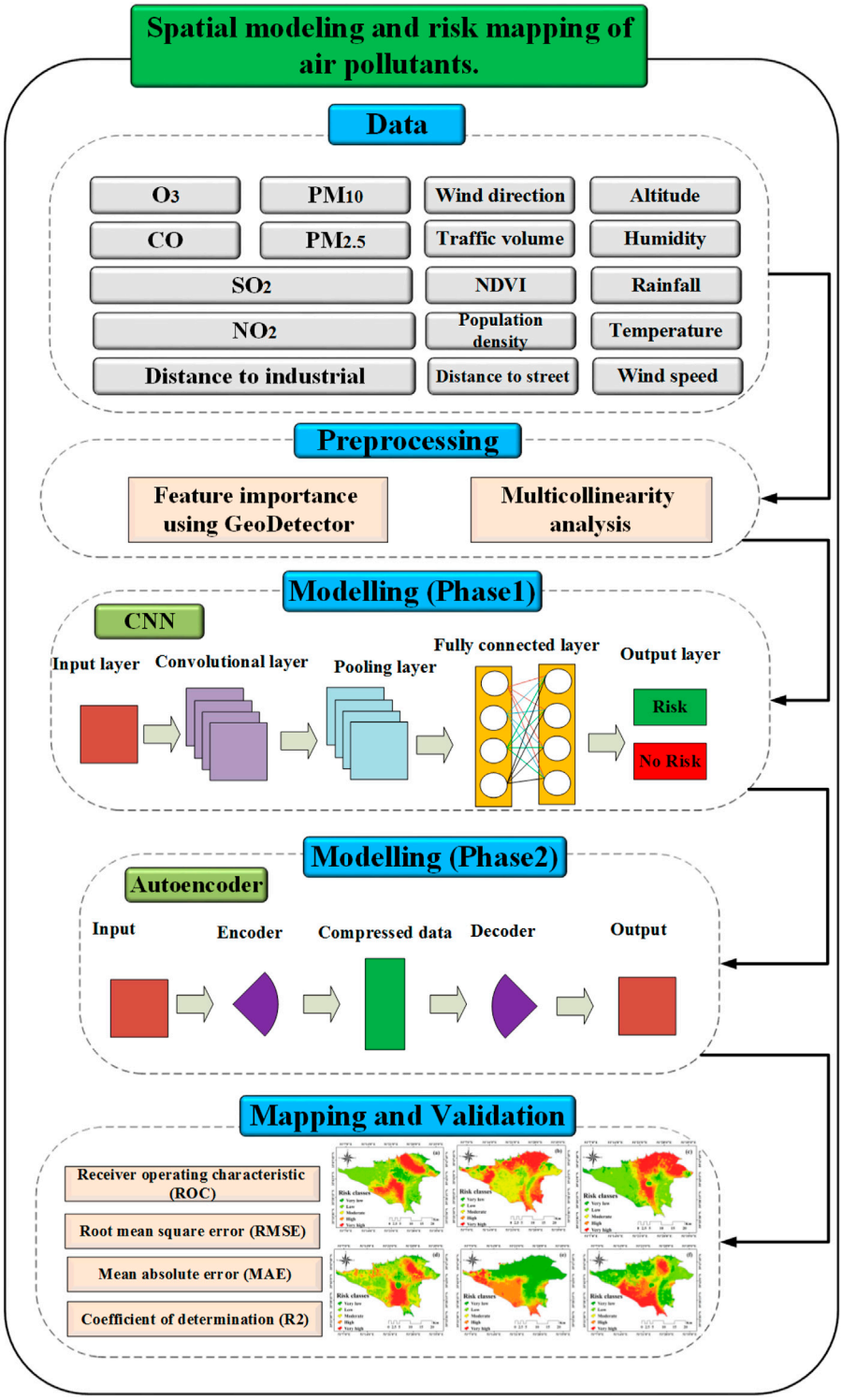


FIGURE 1
Methodology for spatial modeling and risk mapping of air pollutants.

of 1032–1832 m above sea level. The city of Tehran as the most populous city in Iran (9,039,000) in the last 2 decades due to reasons such as the unsustainable development of industrialization and urbanization, the ever-increasing growth of the transport fleet, and the emission of their pollutants, ineffective national environmental air quality standards, and dust storms. The Middle East has faced severe air pollution, especially (PM_{10}, $PM_{2.5}$, O_3, NO_2, SO_2, and CO) (Yousefian et al., 2020). In general, 20% of Iran's energy is consumed in Tehran. The mountain ranges surrounding the city of Tehran stop the flow of humid wind to the capital, so in winter, the cold weather and lack of wind cause the polluted air to be trapped inside the city (Naddafi

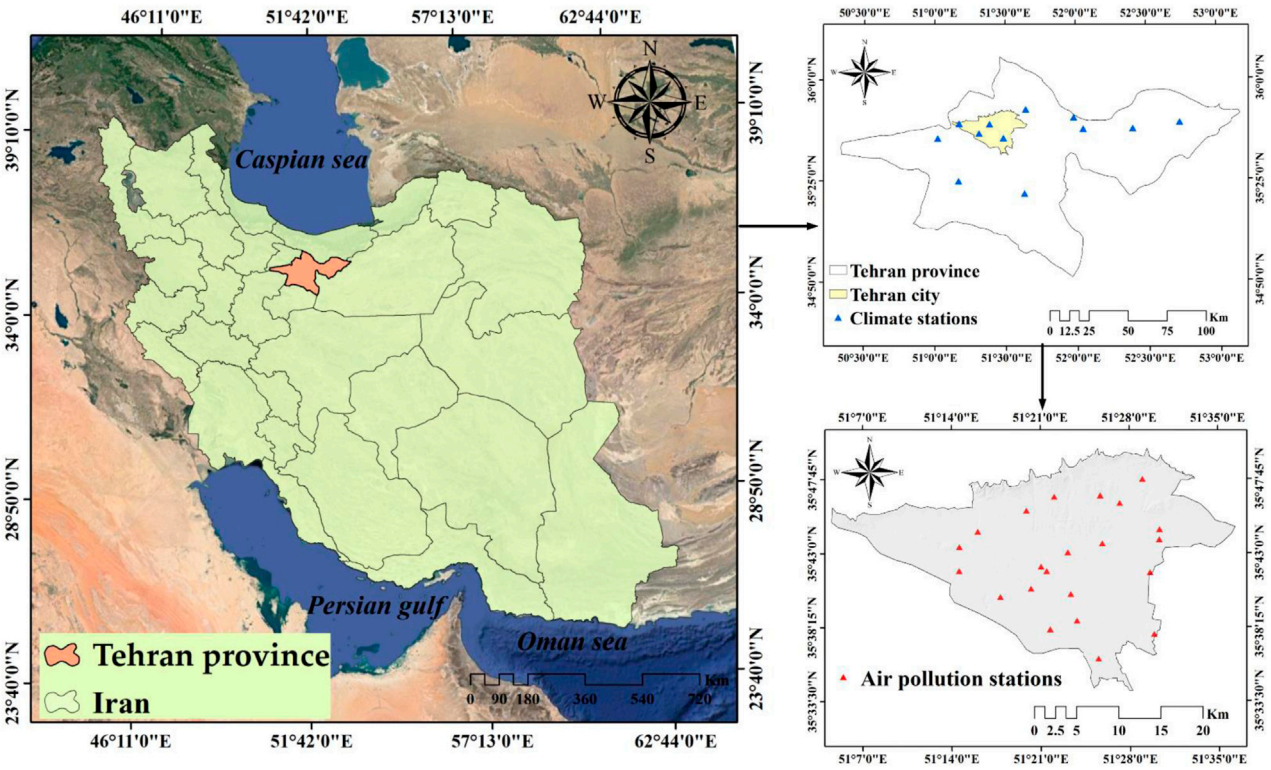


FIGURE 2
Study area with air pollution and climate stations.

et al., 2012). Urban space structures are deeply connected with the urban transportation system (Rodrigue, 2020). The statistics of Tehran indicate the high rate of land consumption in this city, which has caused a high growth in the area and size of the town, which has caused an increase in the distance and the amount of transportation (private cars and public transportation) to carry out administrative and educational activities, and entertainment, in Tehran. Figure 2 displays the geographic location of the study area in Tehran province, Iran, highlighting the air quality control monitoring stations and meteorological stations.

TABLE 1 Characteristics of the six air pollutants.

Parameter	Unit	Min	Max	Mean	Standard Division
NO_2	ppb	38	71	50.08	8.96
PM_{10}	μg/m^3	29	104	74.08	18.39
CO	ppm	1	4	2.26	0.67
$PM_{2.5}$	μg/m^3	21	45	31.6	6.4
O_3	ppb	16	27	20.52	2.76
SO_2	ppb	5	18	9.13	3.24

2.3 Air pollutants

Air pollutants can be categorized as either natural or anthropogenic and can be classified as primary or secondary. Primary pollutants are released directly into the atmosphere from a particular source, retaining the same composition. On the other hand, secondary pollutants are not directly released into the atmosphere and are formed in case of a reaction or interaction of primary pollutants or become another compound in the atmosphere, such as photochemical smog (Bhargav, 2020). Six pollutants, $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3, and CO, are considered for this research. Air pollution is a challenging environmental issue that endangers the health and wellbeing of people worldwide, comprising a complex blend of gaseous components and suspended particles (Özbay, 2012; Bergstra et al., 2018). Air quality in cities in developing countries has gradually deteriorated due to rapid urbanization, population growth, and industrialization (Turalıoğlu et al., 2005). The annual average concentration of air pollutants in Tehran was measured from 1 January 2012, to 1 January 2022, using data collected from 23 air quality monitoring stations located in the city. The characteristics of the six pollutants are shown in Table 1, and the trends of the data in the years 2012–2022 are shown in Figure 3. Maps related to the concentration of pollutants were prepared using kriging interpolation in ArcGIS 10.8 software with a pixel size of 30 × 30 m (Figure 4). For modeling, high-risk areas for each pollutant were converted into occurrence points (with a target value of 1) and the low-risk regions into non-occurrence points (with a target value of 0). In the following, each of these parameters will be explained.

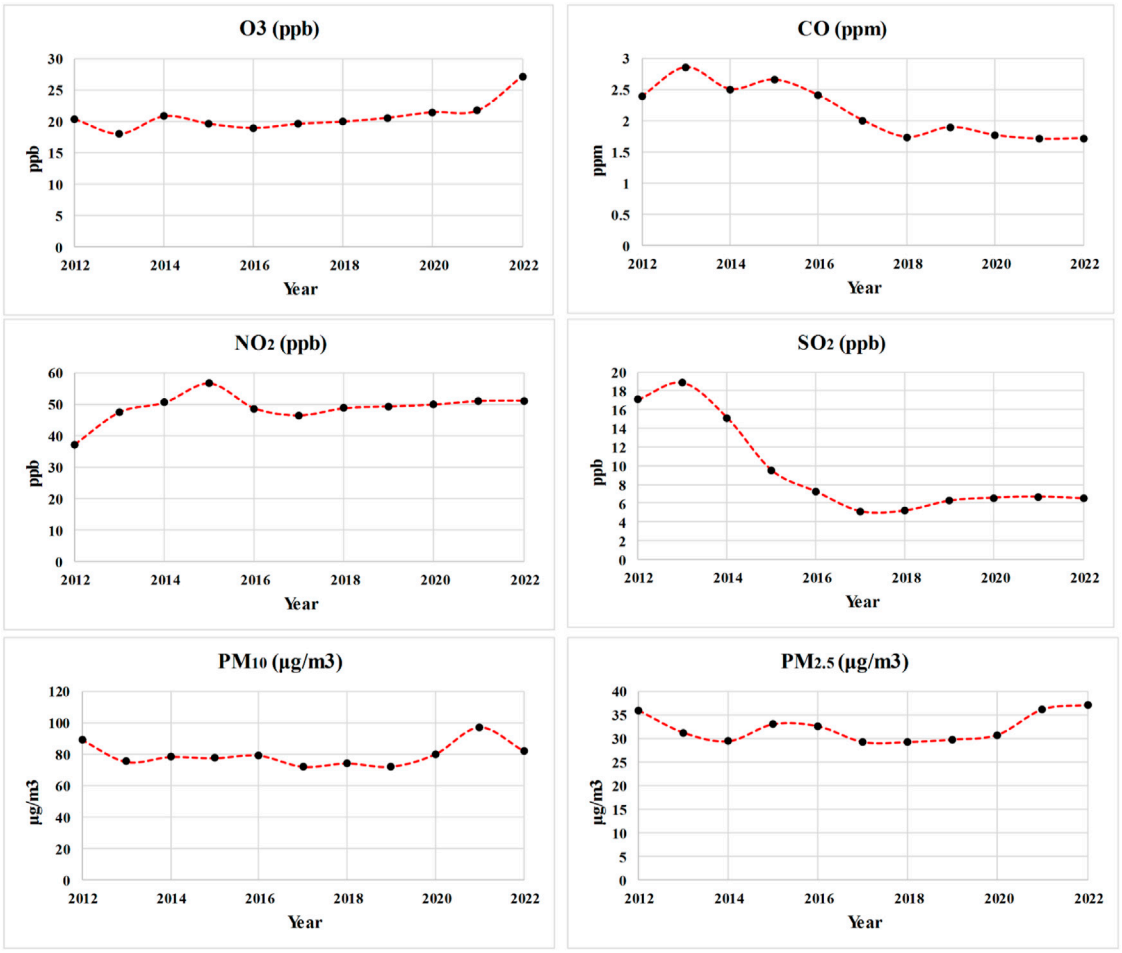


FIGURE 3
The trend of air pollutant concentrations from 2012 to 2022.

➢ Particulate Matter (PM)

Suspended particles include large particles (PM_{10}) and fine particles ($PM_{2.5}$), associated with lung cancer and asthma. PM_{10} can settle in the bronchi and lungs, and $PM_{2.5}$ is the most minor and most dangerous type of suspended particle and can penetrate deep into the respiratory system (Quercia et al., 2015). $PM_{2.5}$ particles are mainly caused by fuel combustion, construction dust, and vehicle exhaust, which cause dust-haze. All types of manufactured combustion and some industrial processes are among the most common human sources of PM_{10} (Özbay, 2012).

➢ Sulfur Dioxide (SO_2)

SO_2 is mainly obtained through the combustion of fossil fuels, biomass burning, and melting ores containing sulfur (Santosa et al., 2008). It is also released through industrial activities and is considered among the harmful gases that affect human, animal, and plant life (Manisalidis et al., 2020). The release of SO_2 in industrial regions can result in serious health concerns such as respiratory irritation, bronchitis, mucus production, bronchospasm, skin redness, eye and mucous membrane damage, and deterioration of cardiovascular health (Chen et al., 2007). Moreover, the environmental consequences of SO_2 include acid rain and soil acidification (Manisalidis et al., 2020).

➢ Nitrogen Oxide (NO_2)

NO_2 is a common traffic-related pollutant that originates from automobile engines, and it is one of the most prevalent air pollutants found in urban regions (Dragomir et al., 2015; Richmond-Bryant et al., 2017). NO_2 is one of the compounds that lead to adverse effects on the environment and human health (Mavroidis and Ilia, 2012), disrupting the sense of smell, burning eyes, throat, and nose, reducing visibility and changing the color of the fabric (Chen et al., 2007).

➢ Carbon Monoxide (CO)

CO is produced due to incomplete or inefficient fuel combustion and affects the blood oxygen transfer in the body and heart (Quercia et al., 2015). CO gas emission and production sources encompass all combustion sources, including motor vehicles, power plants, waste incinerators, domestic gas boilers, and cookers (Vakkilainen, 2017; Manisalidis et al., 2020).

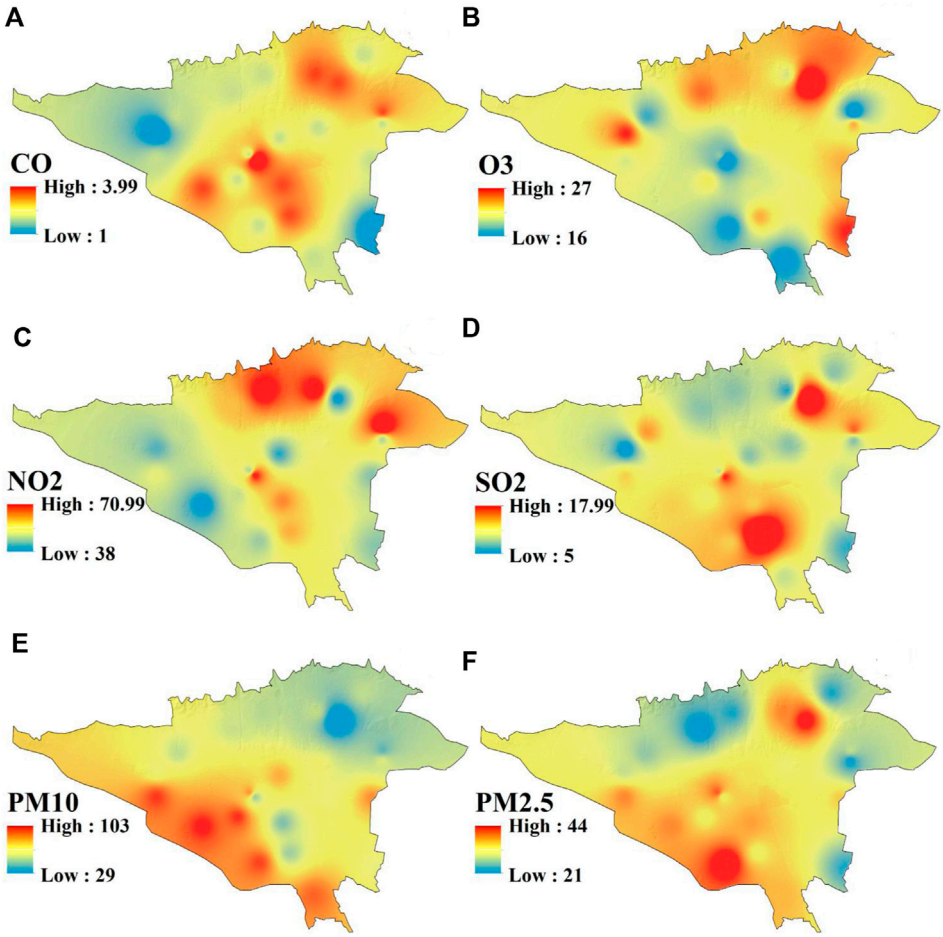


FIGURE 4
Air pollutant concentration maps: **(A)** CO, **(B)** O_3, **(C)** NO_2, **(D)** SO_2, **(E)** PM_{10}, and **(F)** $PM_{2.5}$.

➢ Ozone (O_3)

O_3 is a secondary photochemical pollutant resulting from the oxidation of volatile organic compounds, including benzene, in nitrogen oxides. This colorless, pungent, and reactive gas is the primary component of smog, which is mainly attributed to automobile emissions in urban regions. The concentration of O_3 in urban areas typically increases in the morning, reaches its peak in the afternoon, and decreases at night (Yerramilli et al., 2011).

2.4 Effective factors

The influential factors in this research include meteorological data (rainfall, temperature, humidity, wind direction, and wind speed), altitude, NDVI, distance from street, distance from industrial areas, traffic volume, and population density (Delavar et al., 2019; Shogrkhodaei et al., 2021). Each of the mentioned factors was prepared with a 30 × 30 m pixel size in ArcGIS 10.8 software (Figure 5). In the following, each practical criterion for air pollution has been examined.

- Meteorology data

Air pollution has two natural and human causes, natural causes include volcanic eruptions and severe drought, and human activities include motor vehicle emissions, industry, and the burning of agricultural lands and forests, which cause the release of various types of pollutants with multiple characteristics and effects (Sakti et al., 2023). Pollutants in the atmosphere can be dispersed or diluted under various meteorological conditions, such as rainfall, air temperature, and wind speed (Özbay, 2012). Meteorological data, such as wind speed, wind direction, precipitation, temperature, and humidity, were collected from the National Meteorological Organization. The data were obtained from 12 stations and represented the annual average between 2012 and 2022. The kriging interpolation technique was used in ArcGIS 10.8 to create these maps, with a pixel size of 30 * 30 m. The following discusses the impact of meteorological parameters on air pollutants.

➢ Rainfall

Rainfall is one of the main factors of meteorological conditions that affect air quality and has a specific inhibitory effect on air

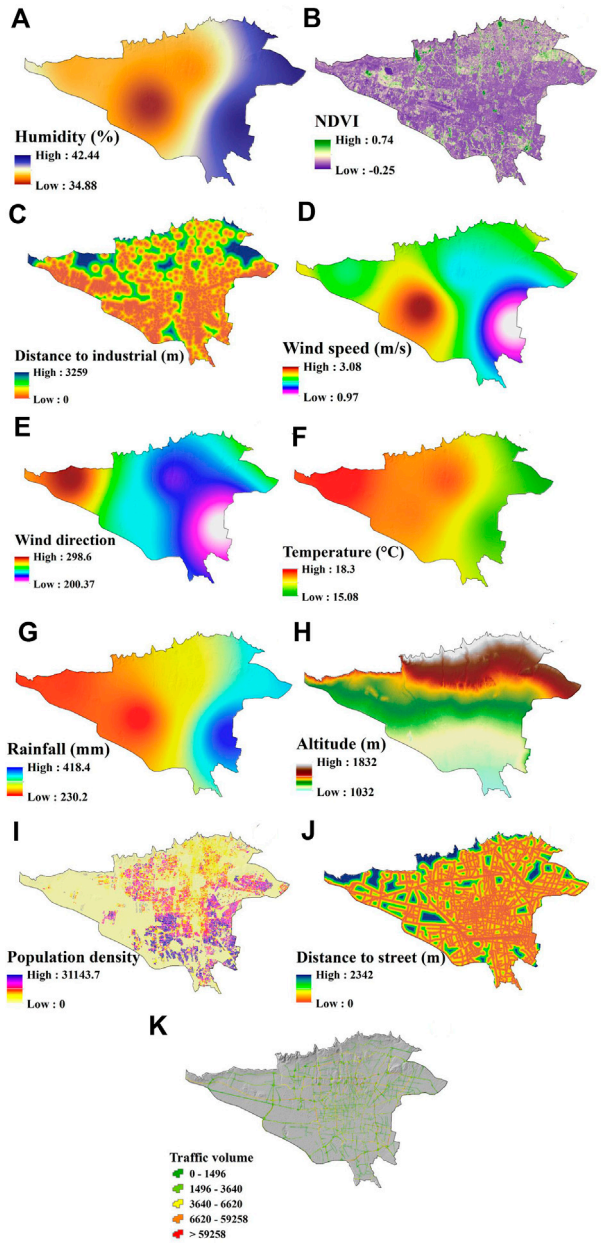


FIGURE 5
Factors affecting air pollution levels: **(A)** Humidity, **(B)** NDVI, **(C)** Distance to industrial, **(D)** Wind speed, **(E)** Wind direction, **(F)** Temperature, **(G)** Rainfall, **(H)** Altitude, **(I)** Population density, **(J)** Distance to street, and **(K)** Traffic volume.

pollutants (Guo and Jiang, 2020; Shukla et al., 2008). Rainfall can affect the concentration of air pollutants by removing gaseous pollution and deposition of suspended particles through atmospheric chemical processes (Kayes et al., 2019).

➢ Temperature

Temperature plays a crucial role in urban air quality, directly influencing gas properties, heterogeneous chemical reaction rates, and the gas-to-particle partitioning process (Aw and Kleeman, 2003). Sunlight and high temperatures stimulate chemical reactions in pollutants and increase smog. The effect of temperature on air pollutants is such that an increase in temperature increases the dispersion and decreases the concentration of contaminants (Shogrkhodaei et al., 2021).

➢ Humidity

Humidity, as one of the meteorological parameters, plays an essential role in air pollutants (their concentration and dispersion) in the urban environment (Endeshaw and Endeshaw, 2020). Most pollutants negatively correlate with relative humidity, so the amount of air pollutants decreases with the increase in humidity (Kayes et al., 2019).

➢ Wind speed and direction

Air quality is affected by wind speed. One case is that wind speed reduces the concentration of pollutants and dilutes them (in areas with higher concentrations). Another issue is that wind speed leads to the entry of contaminants from further distances and increases the concentration of pollutants in an area with a lower concentration (Oleniacz et al., 2016).

• NDVI

Urban vegetation impacts air quality by affecting the sedimentation and dispersion of pollutants (Janhäll, 2015). Urban trees and vegetation are considered an ecosystem regulating service in removing air pollutants (Setälä et al., 2013). The NDVI is a primary indicator of the physiological properties of land vegetation. The NDVI standard was prepared using Landsat-8 satellite images in Google Earth Engine (GEE) (https://earthengine.google.com/) as an annual average from 2012 to 2022. NDVI index was calculated using Eq. 1.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (1)$$

In this equation, the symbol NIR denotes the reflectance in the near-infrared band, and the symbol Red represents the reflectance in the red band. By taking the difference and sum of these reflectance values, the NDVI equation normalizes the values and produces an index that ranges from −1 to +1.

•Altitude

Air pollution is affected by the change in altitude, so the increase in altitude causes an increase in sunlight and causes the problem of photochemical smog (U.S. EPA, 1978). This research prepared the height through a digital elevation model (DEM) with a pixel size of 30 × 30 m through the SRTM (Shuttle Radar Topography Mission) image in GEE platform.

• Distance to industrial areas

Industrial sources located inside and close to city borders are among the influential primary factors of urban air pollution (Hosseini and Shahbazi, 2016). Heavy industry causes the release of many dangerous pollutants in the air that affect health (Bergstra et al., 2018). Industrial areas were extracted through land use layers of industrial areas with a scale of 1:10,000. Subsequently, the

aforementioned criterion was transformed into a raster map with a pixel resolution of 30 × 30 m by employing the Euclidean distance function in ArcGIS 10.8.

•Distance to street

Motor vehicles produce more air pollutants than any other human activity. Motor vehicle emissions from roads can be considered as a mobile line source with an emission rate per unit of road length (Oji and Adamu, 2020). Therefore, the distance from the measurement location to the roads affects the air quality near the streets (Dragomir et al., 2015). The data relating to the roads of Tehran was extracted through the open street map (OSM) (https://www.openstreetmap.org) with a scale of 1:100,000 in 2022. Subsequently, the mentioned criterion was transformed into a raster map with a pixel resolution of 30 × 30 m by utilizing the Euclidean distance function within the ArcGIS 10.8 software.

- Traffic volume

Urban air pollution is primarily caused by traffic emissions (Guarnieri and Balmes, 2014). Monitoring data about pollution near roads shows hot pollution spots in high-traffic areas (Samet, 2007). Traffic congestion increases vehicle emissions, decreases air quality, and increases air pollutants, including CO, CO_2, nitrogen oxides (NO_x), and PM, which cause complications such as death to drivers, passengers, and people who live near the main roads (Zhang and Batterman, 2013). Data on the traffic volume in Tehran city were obtained from the Tehran Traffic Control Company. The data represent the average traffic volume between 2015 and 2020.

- Population density

The expansion of urban areas and population growth has a significant adverse effect on ambient air quality (Kumar et al., 2016), as the rise in population is linked to an increase in the number of vehicles (traffic density) and industrial and commercial operations (Shogrkhodaei et al., 2021). This factor was obtained based on the data from Iran Statistics Center in 2017.

2.5 Methods

2.5.1 Multicollinearity analysis

The problem of multicollinearity arises due to a correlation (strong relationship) between predictors and their lack of independence in a data set. In the models derived from these data, if multicollinearity is not checked, it may lead to wrong analyzes (Garg and Tai, 2013). Variance Inflation Factor (VIF) is a method used to identify multicollinearity in a regression model (Kim, 2019), and a VIF more significant than 10 indicates the presence of multicollinearity (Chen et al., 2018 Eq. 2).

$$\mathrm{VIF} = \frac{1}{\mathrm{Tolerance}} = \frac{1}{1 - \mathrm{R}^2} \tag{2}$$

In the abovementioned equation, the symbol tolerance represents tolerance, and R^2 is the R-squared value of the regression.

2.5.2 Feature importance using GeoDetector method

GeoDetector is a method used to identify and exploit geographic differences and determines the number of driving forces, influencing factors, and multi-factor interactions (Wang and Xu, 2017). This method does not include complex parameter-setting procedures, nor is it limited to the assumptions of classical linear statistical techniques. In this method, if an independent variable significantly affects another independent variable, it will show spatial distribution (Zhang et al., 2022). GeoDetector has four distinct functions: agent detection, interaction detection, hazard detection, and ecological detection (Wang and Xu, 2017). A factor detector is used to detect the spatial heterogeneity of the dependent variable Y and to evaluate the explanatory ability of the independent variable X on Y. The factor detector assesses the effectiveness of the derived q value in capturing the relationship between the variables. The q values obtained from GeoDetector allow the measurement of spatial variations and factor analysis (Jia et al., 2021). The value of q_x was obtained from Eqs 3–5.

$$\mathrm{q_x} = 1 - \frac{\mathrm{SSW}}{\mathrm{SST}} \tag{3}$$

$$SSW = \sum_{h=1}^{l} N_h \sigma_h{}^2 \tag{4}$$

$$\mathrm{SST} = \mathrm{N}\sigma^2 \tag{5}$$

Where SSW denotes the sum of the local variance, while SST represents the global variance. The variable h stands for the number of independent variable categories, N_h and N denote the number of units in zone h and the entire area, respectively. The variable $\sigma_h{}^2$ represents the variance of Y in zone h, and σ^2 denotes the global variance of Y in the entire region.

2.5.3 Convolutional neural network (CNN) algorithm

The traditional ANN in the analysis of complex networks faced the challenge of slowing down the learning process, which Bengio proposed to overcome by a CNN, a neural network that finds local connections between layers (Lu et al., 2017). CNN has achieved remarkable results in various areas of pattern recognition and is particularly useful in reducing the number of parameters in ANN (Albawi et al., 2017). CNN is one of the most widely used deep learning algorithms suitable for spatial data analysis (Khosravi et al., 2020). A CNN architecture generally consists of convolutional, pooling, and fully connected layers (like standard layers in ANN) (Pham et al., 2020). The following is a description of the structure of each layer (Ajit et al., 2020):

- The convolutional layer is the most basic and essential in CNN architecture. This layer performs convolution or multiplication operations on the pixel matrix generated for the target image, resulting in an activation map for that image. The activation map stores all the image's unique features and helps reduce the amount of processed data, which is one of its primary benefits.
- Pooling is a crucial layer that helps reduce the activation map's dimensions while preserving essential features and decreasing

spatial invariance. By reducing the number of learnable features, this layer addresses the issue of overfitting. Pooling also enables CNN to combine all dimensions of an image, allowing it to correctly identify the desired object even if its shape is not in the correct position.

➢ The final layer in the neural network is the fully connected layer, which receives input from the previous layers. All the computations and reasoning are performed in this layer of the data.

2.5.4 Autoencoder (AE) algorithm

AE are neural networks that automatically learn useful features and representations from data (Pinaya et al., 2020). AE is also an unsupervised approach to the neural network method. It does not require data labeling with an operational logic that trains input vectors to be reconstructed as output vectors (Sewani and Kashef, 2020). AE can be used for dimensionality reduction, denoising data, generative modeling, and pre-training deep learning neural networks (Pinaya et al., 2020). An encoder and a decoder make AE architecture. An AE layer has an encoder and a decoder according to Eqs 6, 7, respectively (Zavrak and İskefiyeli, 2020).

$$h = \sigma\left(W_{xh}x + b_{xh}\right) \tag{6}$$

$$z = \sigma\left(w_{hx}x + b_{hx}\right) \tag{7}$$

In the given equations, b and W are referred to as the bias and weight of the neural network, respectively. The symbol σ represents a non-linear transformation function.

2.5.5 Implementing models

The integrated model was implemented using Python in Google Colab, a cloud-based Python development environment. The input data underwent normalization between zero and one to ensure consistent scaling across the different spatial features. This normalization step helps improve the training efficiency and convergence of the model. The experiments and analyses were conducted on a Windows 10 desktop PC with an Intel i7 processor and 16 GB of RAM. The input data was divided into training and testing sets using a 70–30 split. The training set, comprising 70% of the data, was used for model training and parameter optimization. The remaining 30% of the data was reserved for testing the trained model's performance and evaluating its predictive accuracy. In this research, our objective centered on regression and prediction tasks, for which we employed a 1D CNN model architecture. The implemented CNN model was configured with the following parameters: kernel size set to 3, activation function using ReLU, optimizer utilizing Adam, loss function defined as mean squared error (MSE), an epoch count of 400, batch size set to 16, and verbosity level configured to 2 for detailed logging during training.

2.5.6 Validation methods

To extend the model's applicability to unfamiliar outputs, it is necessary to assess its performance by comparing the predicted outcomes from each model with the actual results (Mombeini and Yazdani-Chamzini, 2015). In this study, various indicators such as RMSE, MAE, R^2, and ROC-AUC are employed to evaluate the effectiveness of the model's construction.

➢ RMSE and MAE

RMSE and MAE are indicators that calculate the error between the actual and predicted values (Farahani et al., 2022). The primary difference between MAE and RMSE indices is that MAE assigns equal weight to all errors. Conversely, RMSE penalizes variance by giving more weight to errors with larger absolute values than errors with smaller values (Chai and Draxler, 2014). RMSE and MAE were calculated according to Eqs 8, 9.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(A_i - P_i\right)^2}{N}} \tag{8}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}\left|\left(A_i - P_i\right)\right|}{N} \tag{9}$$

In the above equations, A_i represents the observed value, P_i represents the predicted value, and N is the number of samples.

➢ R^2

R^2 is the variance ratio in the dependent variable that the independent variables can explain (An et al., 2020; Chicco et al., 2021). R^2 was calculated according to Eq. 10.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(A_i - P_i\right)^2}{\sum_{i=1}^{n}\left(A_i - \overline{A_i}\right)^2} \tag{10}$$

In this equation, A_i is the observed value, P_i is the predicted value, and $\overline{A_i}$ is the average of the observed set.

➢ ROC curve

The ROC is a prominent method for evaluating spatial models and a standard tool for determining the accuracy of output maps (Shogrkhodaei et al., 2021). The ROC curve plots the false positive rate (FPR) on the *x*-axis (Eq. 11) against the true positive rate (TPR) on the *y*-axis (Eq. 12) to measure the area under the curve (AUC) as the true-false thresholds change (Pham et al., 2020).

$$x = 1 - \frac{TN}{FP + TN} \tag{11}$$

$$y = \frac{TP}{FN + TP} \tag{12}$$

In this equation, the four data categories in the confusion matrix are TN (True Negative), TP (True Positive), FN (False Negative), and FP (False Positive) (Davis and Goadrich, 2006). AUC is between 0 and 1 (Farahani et al., 2022).

3 Results

3.1 Result of multicollinearity test

A multicollinearity test was performed to assess the presence of multicollinearity among the independent variables utilized in the geographic modeling and risk mapping of the six air contaminants. The results of the test, presented in Table 2, indicate the levels of VIF for each independent variable. From the results, it can be observed that none of the independent variables have VIF values exceeding

TABLE 2 Multicollinearity test results on factors affecting air pollution.

Independent variables	$PM_{2.5}$	PM_{10}	SO_2	NO_2	CO	O_3
Altitude	1.8	1.77	2.09	1.6	1.5	1.5
Humidity	4.4	4.3	5.3	5.4	4.8	4.4
Distance to industrial	1.2	1.3	1.2	1.2	1.2	1.2
NDVI	1.1	1.12	1.14	1.07	1.16	1.15
Population density	1.09	1.16	1.17	1.14	1.2	1.2
Rainfall	5.3	5.2	6.2	5.9	5.6	5.5
Distance to street	1.13	1.13	1.14	1.19	1.2	1.2
Temperature	5.9	3.6	4.9	5.2	5.1	5.2
Traffic volume	1.2	1.09	1.2	1.13	1.15	1.09
Wind direction	3.9	3.4	4.1	3.6	3.5	3.6
Wind speed	6.7	6.8	6.3	6.5	6.6	6.3

the threshold limit of 10. This indicates no severe multicollinearity among the independent variables, suggesting that they can be considered individually and collectively in the spatial modeling and risk mapping analysis. The VIF values, ranging from 1.07 to 6.80, indicate that the independent variables have relatively low to moderate levels of correlation with each other. This suggests that the variables provide unique information and do not excessively duplicate each other's predictive power.

3.2 Result of feature importance

The GeoDetector method was employed to determine the importance of different parameters on air pollutants (Figure 6). The analysis revealed distinct findings for each pollutant. CO, temperature, wind speed, and wind direction emerged as the most significant parameters. Variations notably influenced the levels of CO in these factors. In the case of O_3, humidity, precipitation, and altitude were identified as the primary criteria affecting its concentration. Altitude plays a crucial role in the formation and distribution of ozone in the atmosphere.

Conversely, for PM_{10}, altitude, wind direction, and wind speed were deemed the most influential parameters. These factors influenced the dispersion and transport of PM_{10} particles. Regarding NO_2, altitude, rainfall, and wind direction were found to have the most significant impact. Altitude affected the vertical distribution of NO_2, while rainfall and wind direction influenced its dispersion and movement. Similarly, for $PM_{2.5}$, altitude, rainfall, and temperature were identified as the key parameters. Altitude affected the vertical distribution of $PM_{2.5}$ particles, while rainfall and temperature were crucial in their formation and dispersion. Lastly, for SO_2, temperature, wind direction, and altitude were determined as the most important parameters. Temperature played a role in the chemical reactions involving SO_2, while wind direction and altitude affected its transport and dispersion. In general, altitude, wind direction, wind speed, rainfall, and temperature parameters had the most significant effect on pollutants in the study area.

3.3 Model development

The AE, comprising encoder and decoder layers, is a pre-training step to learn a compact and efficient representation of the input data. The CNN, however, is designed to capture spatial patterns and dependencies within the pollutant data. The input data includes various spatial features such as altitude, humidity, distance to industrial areas, NDVI, population density, rainfall, distance to the street, temperature, traffic volume, wind direction, and wind speed. The Autoencoder's encoded features serve as input to the CNN, which then extracts spatial features. The model's weights, biases, learning rates, regularization techniques, and dropout rates are randomly initialized and updated during the training process using the Adam optimizer. A loss function is utilized to measure the difference between the predicted pollutant concentrations and the actual measurements to assess the model's performance. Common regression loss functions, such as mean squared error (MSE), are commonly used. The results of the loss functions for all pollutants, as shown in Figure 7, indicate the convergence and effectiveness of the integrated CNN-AE model. The loss function values for the training and test data decrease throughout training, demonstrating the model's ability to learn and capture the underlying patterns in the pollutant data. The decreasing trend of the loss function values suggests the model successfully minimizes the discrepancy between the predicted pollutant concentrations and the actual measurements during training. This indicates that the model is learning to make accurate predictions and is effectively capturing the complex relationships within the data. The decreasing loss function values in the training and test data support the notion that the integrated CNN-AE model successfully learns and generalizes to unseen data, highlighting its ability to capture the spatial patterns and dependencies of the air pollutants.

Additionally, metrics such as MAE, RMSE, and R^2 are calculated to assess the accuracy and predictive power of the model (Table 3; Table 4). For the pollutant $PM_{2.5}$, the CNN model exhibited reasonably good performance, achieving an R2 of 0.889. The corresponding RMSE and MAE values were 0.166 and 0.046,

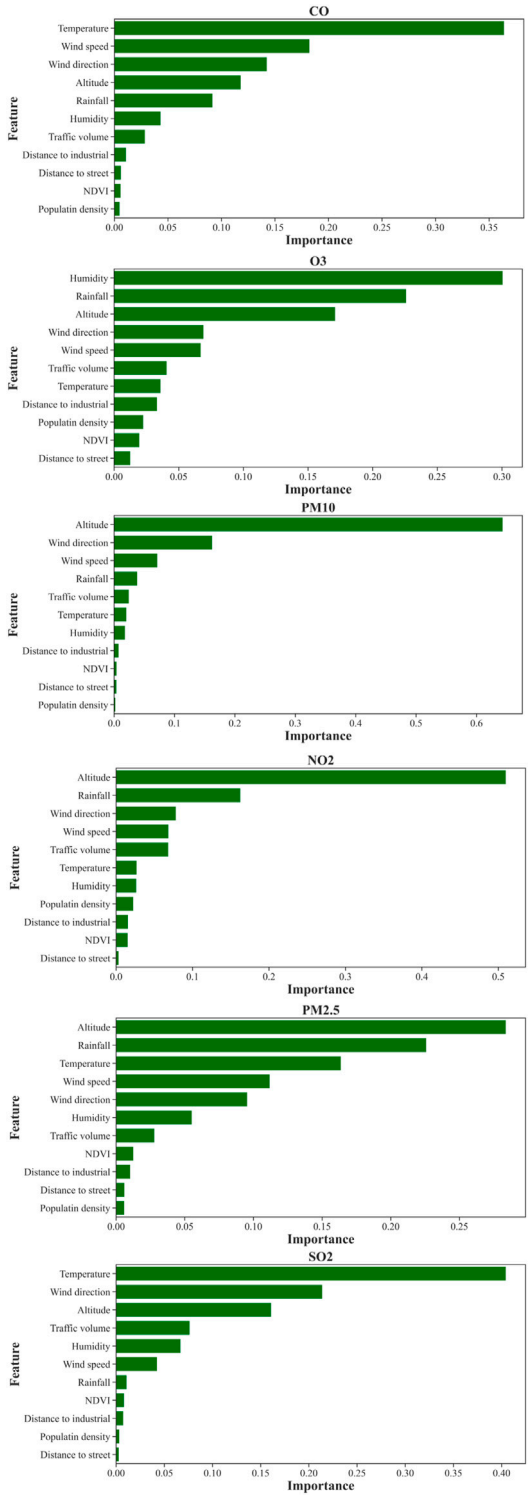


FIGURE 6
Feature importance results using GeoDetector.

respectively. However, the CNN-AE model surpassed the CNN model's performance, demonstrating an improved R^2 of 0.969. Moreover, the RMSE and MAE values for the CNN-AE model were 0.087 and 0.157, respectively, indicating better accuracy and precision in predicting $PM_{2.5}$ concentrations.

Regarding the pollutant PM_{10}, both models performed exceptionally well. The CNN model achieved an impressive R^2 of 0.972, suggesting that the model could explain approximately 97.2% of the PM_{10} concentration variance. Additionally, the CNN model exhibited low RMSE and MAE values of 0.082 and 0.053, respectively. The CNN-AE model further enhanced the prediction accuracy, yielding an even higher R^2 of 0.98. The RMSE and MAE values for the CNN-AE model were 0.0701 and 0.045, respectively, indicating a significant improvement over the CNN model. For the pollutant SO_2, both the CNN and CNN-AE models demonstrated commendable performance. The CNN model achieved an R^2 of 0.951, suggesting that the model could explain approximately 95.1% of the SO_2 concentration variability. The corresponding RMSE and MAE values were 0.11 and 0.075, respectively. The CNN-AE model showed similar performance, with an R2 of 0.955, indicating a comparable ability to explain the variability in SO_2 concentrations. The RMSE and MAE values for the CNN-AE model were 0.105 and 0.067, respectively, demonstrating their effectiveness in predicting SO_2 levels.

Regarding the pollutant NO_2, both models exhibited solid predictive capabilities. The CNN model achieved an R^2 of 0.972, indicating that the model could explain approximately 97.2% of the NO_2 concentration variability. The RMSE and MAE values were 0.083 and 0.054, respectively, suggesting accurate predictions. The CNN-AE model outperformed the CNN model, attaining an exceptional R^2 of 0.994. The RMSE and MAE values for the CNN-AE model were significantly lower at 0.038 and 0.032, respectively, indicating superior precision and accuracy in predicting NO_2 concentrations. For the pollutant O_3, both models demonstrated satisfactory performance. The CNN model achieved an R^2 of 0.949, suggesting that the model could explain approximately 94.9% of the O_3 concentration variability. The RMSE and MAE values were 0.112 and 0.08, respectively. The CNN-AE model improved the prediction accuracy with an R^2 of 0.96.

Regarding the CO pollutant, the CNN model demonstrated a high level of performance, as indicated by an R^2 value of 0.952. This suggests that the model's predictions account for around 95.2% of the variability in CO concentrations. The RMSE and MAE values for CO were calculated as 0.109 and 0.073, respectively. Notably, the CNN-AE further enhanced the accuracy of CO predictions. The CNN-AE model achieved an improved R^2 value of 0.978, indicating that the model captured approximately 97.8% of the CO concentration variability. The corresponding RMSE and MAE values were calculated as 0.073 and 0.044, respectively.

Moving on to the test data, the CNN exhibited moderate performance with R^2 values ranging from 0.57 to 0.715 for six pollutants. The RMSE values ranged from 0.265 to 0.324, indicating some difference between the predicted and actual values. The MAE values ranged from 0.162 to 0.233, representing the average absolute difference between predicted and actual values. In contrast, the CNN-AE improved performance on the test data compared to the CNN. It achieved higher R^2 values ranging from 0.681 to 0.829, indicating a better fit. The lower RMSE values, ranging from 0.205 to 0.281, suggested more accurate predictions. The MAE values ranged from 0.106 to 0.185, indicating a more negligible average absolute difference between predicted and actual values compared to the CNN.

In summary, integrating the AE with the CNN algorithm showed promising results in air quality management. The CNN

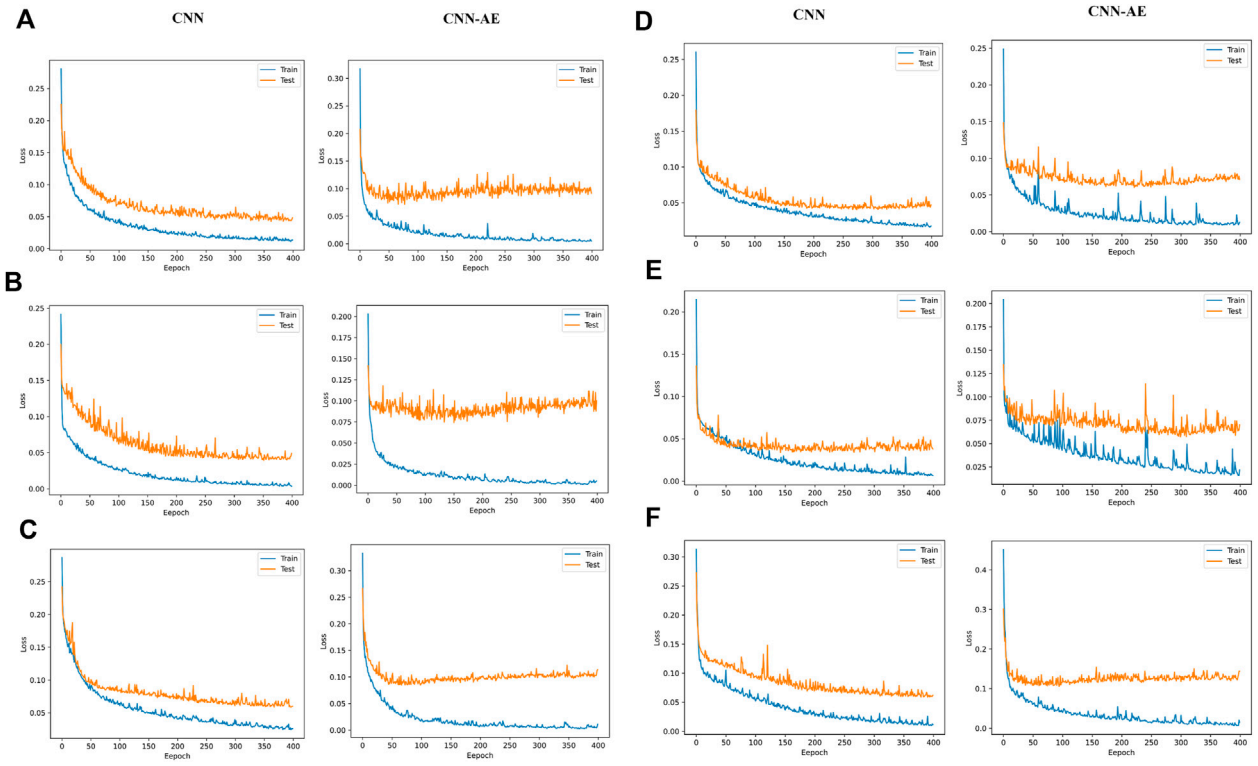


FIGURE 7
Loss function comparison of CNN and CNN-AE models: **(A)** CO, **(B)** O_3, **(C)** NO_2, **(D)** SO_2, **(E)** PM_{10}, and **(F)** $PM_{2.5}$.

TABLE 3 Result of air pollution modeling in the training phase.

Pollutants	CNN			CNN-AE		
	R^2	RMSE	MAE	R^2	RMSE	MAE
$PM_{2.5}$	0.889	0.166	0.046	0.969	0.087	0.157
PM_{10}	0.972	0.082	0.053	0.98	0.0701	0.045
SO_2	0.951	0.11	0.075	0.955	0.105	0.067
NO_2	0.972	0.083	0.054	0.994	0.038	0.032
O_3	0.949	0.112	0.08	0.96	0.099	0.073
CO	0.952	0.109	0.073	0.978	0.073	0.044

TABLE 4 Result of air pollution modeling in the test phase.

Pollutants	CNN			CNN-AE		
	R^2	RMSE	MAE	R^2	RMSE	MAE
$PM_{2.5}$	0.715	0.265	0.162	0.793	0.226	0.157
PM_{10}	0.691	0.277	0.163	0.829	0.205	0.106
SO_2	0.608	0.312	0.222	0.77	0.238	0.164
NO_2	0.678	0.28	0.167	0.763	0.24	0.151
O_3	0.57	0.324	0.233	0.681	0.281	0.185
CO	0.606	0.312	0.178	0.805	0.22	0.145

and CNN-AE models exhibited strong performance in the training phase, with the CNN-AE model consistently outperforming the CNN. Although there was a slight decrease in performance during the testing phase, the CNN-AE model maintained its superiority over CNN. Figure 8 shows the fitting diagram of the training and test data on the target data.

3.4 Creation of risk map and validation

Using the trained model, the CNN-AE model estimated pollutant concentrations for each location in the study area. These estimated concentrations were then assigned risk levels to different regions based on classification criteria. The risk levels could be categorized as very low, low, moderate, high, and very high, representing varying degrees of pollution severity (Figure 9). The risk maps were generated by overlaying the estimated pollutant concentrations onto a geographical map of the study area. Each region was color-coded according to the assigned risk level, providing an intuitive visualization of the pollution hotspots and areas of concern, and according to the risk maps generated from the CNN-AE model, the southwest and northeast regions exhibited higher risk levels for CO pollution. Concerning O_3 pollution, elevated risk levels were observed in the north, east, and west areas. The risk of NO_2 pollution was particularly pronounced in the north and central regions. In the case of SO_2 pollution, the risk

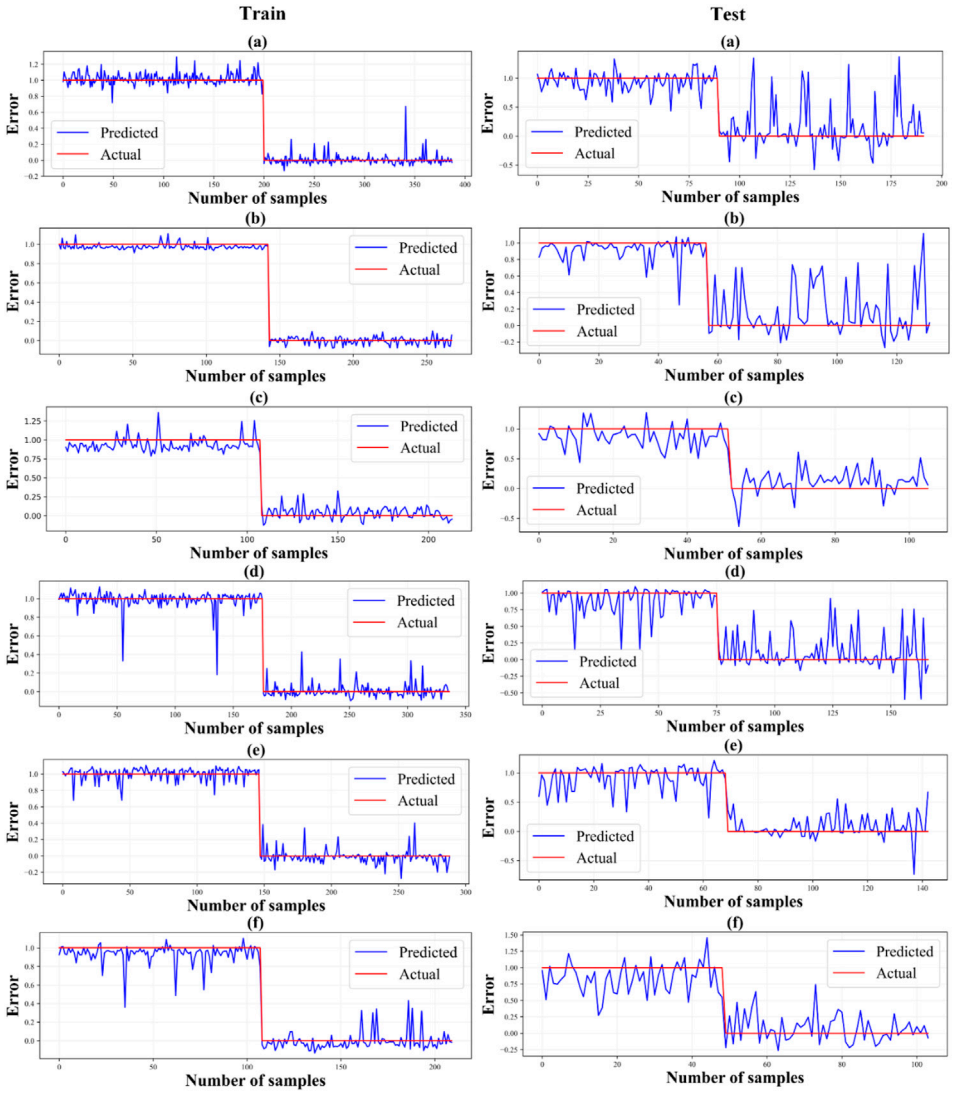


FIGURE 8
Error diagram in training and test data. **(A)** CO, **(B)** O_3, **(C)** NO_2, **(D)** SO_2, **(E)** PM_{10}, and **(F)** $PM_{2.5}$.

was concentrated in the south and northeast areas. PM_{10} pollution posed a higher risk in the west and southwest regions, while $PM_{2.5}$ pollution was more prominent in the southern part.

Several evaluation metrics were employed to assess the effectiveness of the risk maps generated by the CNN-AE method, including the ROC curve, AUC, and Youden index J. These metrics were used to analyze the performance of the risk maps in terms of their ability to accurately discriminate between different risk levels. The evaluation results, as presented in Table 5 and Figure 10. For NO_2, an AUC of 0.964 was obtained, indicating a high level of discrimination between different risk levels. The Youden index J was 0.8936, further confirming the model's ability to identify the optimal threshold for risk classification. The Standard Error was 0.0235, and the 95% Confidence Interval ranged from 0.903 to 0.991, indicating high precision in the risk map. The z statistic value was 19.72, and the significance level was $p < 0.0001$, demonstrating the statistical significance of the results.

Similarly, for PM_{10}, an AUC of 0.95 was achieved, indicating good discriminatory power. The Youden index J was 0.8276, highlighting the model's effectiveness in identifying risk thresholds. The Standard Error was 0.0175, and the 95% Confidence Interval ranged from 0.907 to 0.977, indicating a high confidence level in the risk map. The z-statistic value was 25.764, and the significance level was $p < 0.0001$, further confirming the statistical significance of the findings. The performance of the CNN-AE algorithm was also evaluated for CO, $PM_{2.5}$, O_3, and SO_2. The AUC values for CO, $PM_{2.5}$, and O_3 were 0.896, 0.878, and 0.877, respectively, demonstrating moderate to high discriminatory power. The Youden index values were 0.75, 0.7368, and 0.7292, indicating the model's ability to identify suitable risk thresholds. The Standard Errors were 0.0321, 0.0298, and 0.0391, respectively, showing the precision of the risk maps. The 95% Confidence Intervals ranged from 0.827 to 0.944 for CO, 0.815 to 0.926 for $PM_{2.5}$, and 0.794 to 0.935 for O_3, further strengthening the reliability of the risk estimates. The z-statistic values were 12.332, 12.69, and 9.632, respectively, and the significance levels were $p < 0.0001$ for all three pollutants,

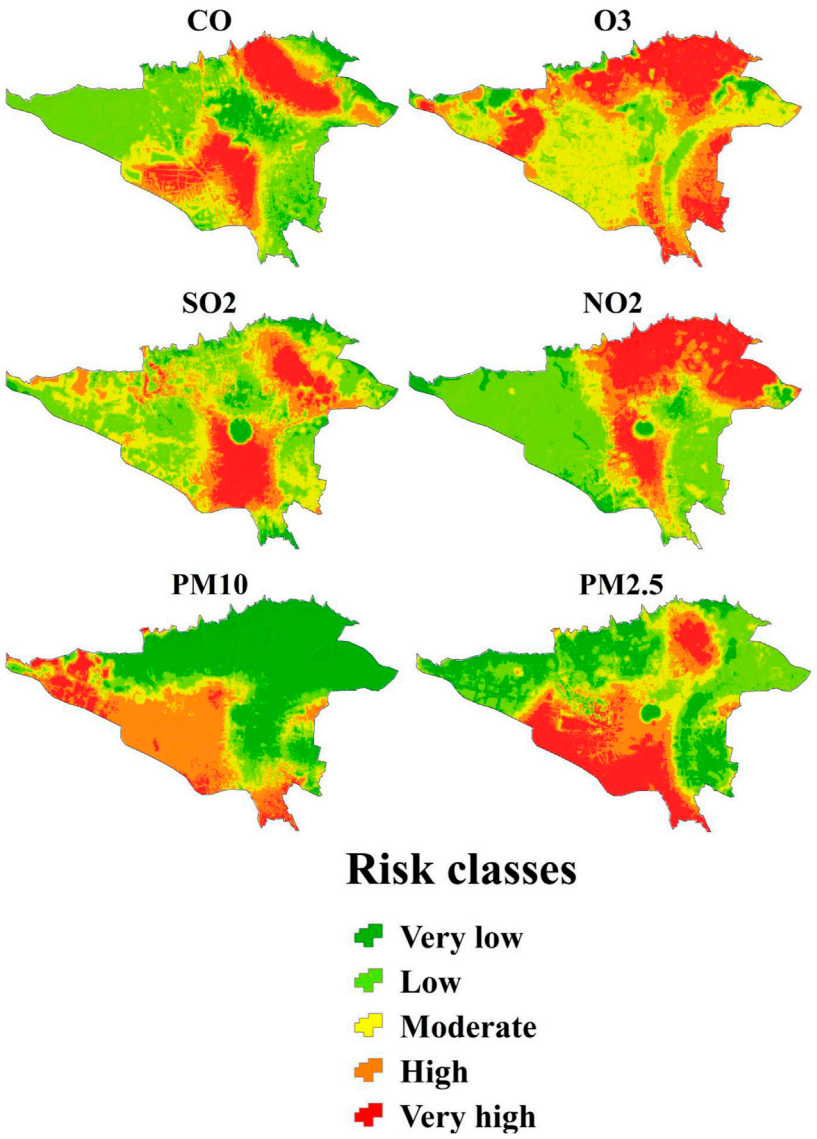


FIGURE 9
Risk map of different pollutants.

TABLE 5 Validation result of air pollutants risk mapping.

Parameter	AUC	Youden index J	Standard error	95% confidence interval	z statistic	Significance level
NO_2	0.964	0.8936	0.0235	0.903 to 0.991	19.72	$p < 0.0001$
PM_{10}	0.95	0.8276	0.0175	0.907 to 0.977	25.764	$p < 0.0001$
CO	0.896	0.7500	0.0321	0.827 to 0.944	12.332	$p < 0.0001$
$PM_{2.5}$	0.878	0.7368	0.0298	0.815 to 0.926	12.690	$p < 0.0001$
O_3	0.877	0.7292	0.0391	0.794 to 0.935	9.632	$p < 0.0001$
SO_2	0.811	0.6308	0.0425	0.733 to 0.874	7.317	$p < 0.0001$

underscoring the statistical significance of the observed results. Lastly, for SO_2, an AUC of 0.811 was obtained, indicating an acceptable level of discrimination between risk levels. The Youden index J was 0.6308, suggesting the model's capability to identify appropriate risk thresholds. The Standard Error was 0.0425, and the 95% Confidence Interval ranged from 0.733 to 0.874, providing

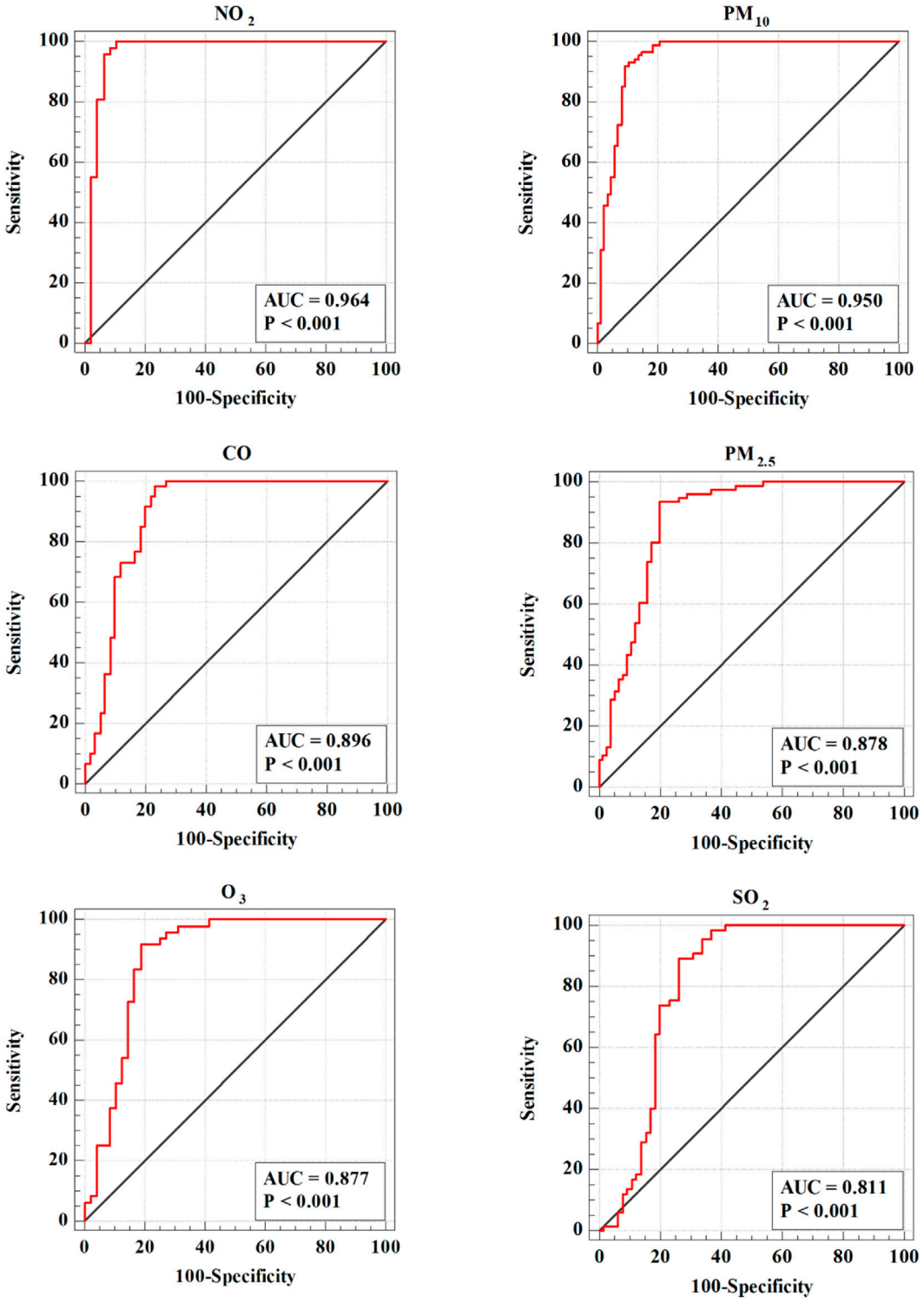


FIGURE 10
Validation of risk maps by ROC curve.

a reliable estimate of the risk map. The z statistic value was 7.317, and the significance level was $p < 0.0001$, affirming the statistical significance of the results. Integrating the AE with the CNN algorithm proved effective in spatial modeling and risk mapping of the six air pollutants. The high AUC values, significant Youden index values, narrow confidence intervals, and low p-values indicate the model's ability to discriminate between different levels of pollutant risk and its statistical reliability. These results contribute to our understanding of the spatial distribution and potential.

4 Discussion

The study's outcomes indicate that combining AE with CNN algorithms is a successful approach for spatial modeling and risk mapping of six air pollutants. By combining the strengths of these two techniques, we overcame-overcame the limitations of traditional modeling approaches and achieved more accurate predictions of air pollutant concentrations. This section discusses the key findings, implications, limitations, and potential future directions of the research. One of the major findings of this study is the significant improvement in modeling accuracy achieved through the CNN-AE fusion approach. Integrating the autoencoder allowed for extracting essential features and patterns from the air pollutant data, effectively reducing dimensionality while preserving relevant information (Dairi et al., 2021).

On the other hand, the CNN leveraged the spatial relationships and patterns in the data, enabling more precise modeling of the pollutant concentrations across the study area (Jiang et al., 2022). As a result, the combined model outperformed traditional modeling approaches, as evidenced by the reduced MAE and RMSE values. The superior performance of the CNN-AE model can be attributed to the benefits provided by the AE component. The AE enables the model to learn a compact and meaningful representation of the input data, which enhances its ability to extract relevant features and patterns. This feature extraction capability is significant in air quality management, as various complex and interrelated factors influence pollutant levels (Cheng et al., 2018; Shankar and Parsana, 2022).

The GeoDetector method assessed the importance of different parameters on various air pollutants, revealing crucial insights for policymakers and researchers. For the CO pollutant, the observed influence of temperature, wind speed, and wind direction can be attributed to their impact on the combustion processes and emissions. Higher temperatures may enhance CO's chemical reactions, increasing pollutant levels (Noyes et al., 2009). Wind speed and direction play a crucial role in the dispersion of CO emissions, affecting the spatial distribution and concentration of the pollutant (Gorai et al., 2015). Regarding the O_3 pollutant, humidity, rainfall, and altitude are important factors. The formation of ozone is primarily influenced by photochemical reactions that occur when nitrogen oxides (NOx) and volatile organic compounds (VOCs) are present in sunlight (Swamy et al., 2012). Humidity and precipitation can influence these reactions by altering the availability of reactants and the rate of chemical transformations (Bell, 2020). Altitude plays a role in determining the amount of solar radiation and the temperature conditions conducive to ozone formation (Zhao et al., 2019). For PM_{10}, altitude, wind direction, and wind speed have significant impacts. Altitude affects the dispersion and transport of PM_{10} particles, with higher altitudes often leading to increased atmospheric mixing and dilution of pollutants (Li et al., 2019). Wind direction and speed determine the pathways and distances PM_{10} particles can travel, influencing their spatial distribution and concentration (Wang et al., 2010).

Regarding the NO_2 pollutant, the altitude parameter indicates the vertical distribution of NO_2 emissions (Salmond et al., 2013). Higher emissions released from industrial sources or vehicle exhausts closer to the ground can contribute to increased levels of NO_2 (Richter et al., 2005). Rainfall can play a role in removing NO_2 from the atmosphere through wet deposition, while wind direction influences the spatial distribution and transport of NO_2 emissions (Matejko et al., 2009). For $PM_{2.5}$, altitude, rainfall, and temperature exhibit notable effects. Altitude influences the vertical distribution of $PM_{2.5}$ particles, with emissions and sources at different heights impacting their ground-level concentration (Peng et al., 2015). Rainfall can remove $PM_{2.5}$ particles from the atmosphere, lowering pollutant levels (Nowak et al., 2013). Temperature can influence the chemical reactions and physical processes involved in forming, transforming, and dispersing $PM_{2.5}$ particles (Su et al., 2020). Finally, for the SO_2 pollutant, temperature affects the rates of chemical reactions involving SO_2. Higher temperatures can facilitate the conversion of SO_2 into other secondary pollutants, such as sulfuric acid aerosols (He et al., 2014). Wind direction and height play a role in the transport and dispersion of SO_2 emissions, influencing the spatial distribution and concentration of the pollutant (Hong et al., 2021).

Our analysis revealed higher risk levels of SO_2 pollution in Tehran's northeastern, central, and southern regions. This heightened risk can be attributed to the concentration of industrial zones and higher population density in these areas. Industrial activities and dense urban settlements are known to be significant sources of SO_2 emissions, contributing to elevated pollution levels. Our findings depicted higher risk levels of $PM_{2.5}$ and PM_{10} pollution in Tehran's southwestern and southern regions. This pattern can be attributed to the concentration of industrial areas in these zones. Industrial activities are a significant source of particulate matter emissions, contributing to higher pollution levels in nearby residential and commercial areas. The risk maps for CO indicated increased risk levels in the southwestern and northeastern parts of Tehran. This observation can be linked to the density of road networks and higher traffic volume in these areas. The combustion of fossil fuels in vehicles releases CO emissions, resulting in elevated concentrations near significant roadways and urban centers. The risk maps for O_3 pollution indicated elevated risk levels in the northern, southern, and eastern parts of Tehran. This heightened risk is associated with increased traffic emissions, NO_x, indirectly contributing to O_3 formation through photochemical reactions. Additionally, Tehran's central, northeastern, and eastern areas exhibited higher NO_2 concentrations due to population density and increased vehicular traffic.

Despite the valuable insights gained from this research on air quality management using spatial modeling, risk mapping, and the integration of the AE with the CNN algorithm, it is important to acknowledge certain limitations and offer suggestions for future research. Firstly, the accuracy of the models heavily relies on the quality and representativeness of the input data. Any inaccuracies or biases in the monitoring data could affect the reliability of the models and risk maps. Additionally, the spatial criteria used in the analysis are based on existing knowledge and assumptions about factors influencing air pollution. There may be other unaccounted spatial parameters that could affect the models' accuracy. Future studies could explore incorporating more comprehensive datasets and advanced feature selection techniques to enhance the modeling accuracy.

Furthermore, the evaluation metrics used in this study, such as MAE and RMSE, provide an overall assessment of the modeling performance. However, it is essential to consider additional evaluation measures, such as spatial validation techniques, to

assess the goodness of fit and the model's ability to capture spatial patterns accurately. This can provide further insights into the reliability and generalizability of the risk maps generated. In terms of future directions, this research opens avenues for exploring additional techniques and methodologies to enhance air quality modeling and risk mapping. For example, incorporating spatiotemporal modeling approaches could capture the dynamic nature of air pollution and improve the accuracy of predictions. Furthermore, integrating other machine learning algorithms or hybrid models could yield even better results by leveraging the strengths of different techniques.

The improved spatial modeling and risk mapping techniques developed in this study provide valuable tools for policymakers and environmental regulators to design targeted interventions and implement evidence-based policies for air quality management. By identifying pollution hotspots and understanding the underlying factors contributing to elevated pollutant levels, policymakers can prioritize resources and implement mitigation measures to reduce exposure and protect public health. Furthermore, integrating advanced modeling techniques can enhance the effectiveness of regulatory initiatives to reduce emissions from industrial facilities, transportation networks, and other pollution sources.

The successful fusion of AE with CNN opens up new avenues for air quality modeling and risk assessment research. Future studies could explore further enhancements to the modeling framework by incorporating additional data sources, refining feature extraction algorithms, and integrating spatiotemporal modeling approaches to capture the dynamic nature of air pollution. Additionally, research efforts could focus on investigating the interactions between different pollutants and identifying synergistic effects on human health, ecosystems, and climate change. Furthermore, interdisciplinary collaborations between researchers from various domains, including environmental science, computer science, and public health, can facilitate the development of innovative solutions to address complex air quality challenges.

5 Conclusion

This research presents a novel and innovative approach for spatial modeling and risk mapping of six air pollutants by combining AE with a CNN algorithm. Integrating these two techniques has significantly improved modeling accuracy and the generation of informative risk maps. The research results indicate that the integrated CNN-AE model outperforms the standalone CNN model regarding predictive accuracy. The evaluation of the models on train and test data further confirmed the superiority of the CNN-AE model, as it achieved higher R^2 values, lower RMSE values, and smaller MAE values than the CNN model. These findings suggest that integrating the AE with the CNN algorithm enhances the model's ability to capture and utilize the spatial relationships in the pollutant data. In the study area, the pollutants were most influenced by specific parameters, namely, altitude, wind direction, wind speed, rainfall, and temperature, as determined by applying the GeoDetector method.

The risk maps generated by the CNN-AE model indicated distinct pollution patterns across different regions. The southwest and northeast regions showed higher risk levels for CO pollution. Elevated risk levels for O_3 pollution were observed in the north, east, and west areas. The north and central regions exhibited a pronounced risk of NO_2 pollution. The risk of SO_2 pollution was concentrated in the south and northeast areas. PM_{10} pollution posed a higher risk in the west and southwest regions, while $PM_{2.5}$ pollution was more prominent in the southern part. The risk maps generated through the integrated methodology provide valuable insights for air quality management. By visualizing the spatial distribution of the pollutant concentrations, these risk maps help identify high-risk areas and pollution hotspots. This information is crucial for policymakers, environmental agencies, and stakeholders to prioritize mitigation efforts and allocate resources effectively. The risk maps can also support decision-making processes, facilitating the development of targeted interventions to reduce pollutant levels and protect public health. For future research, it is suggested that the CNN-AE model be adapted and validated across diverse geographical regions to ensure generalizability and robustness. Incorporating real-time data from sensors and satellite imagery could enhance the model's real-time air quality monitoring applicability. Additionally, expanding the methodology to include a broader range of pollutants and investigating the impact of climate change on pollution patterns will provide comprehensive assessments. Linking the risk maps with health impact assessments could offer valuable insights into public health implications, supporting informed policy development.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

AB: Conceptualization, Data curation, Formal Analysis, Investigation, Resources, Software, Visualization, Writing–original draft. MM: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing–review and editing. MK: Funding acquisition, Methodology, Resources, Validation, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abu El-Magd, S., Soliman, G., Morsy, M., and Kharbish, S. (2022). Environmental hazard assessment and monitoring for air pollution using machine learning and remote sensing. *Int. J. Environ. Sci. Technol.*, 1–14. doi:10.1007/s13762-022-04367-6

Agarwal, S. K. (2021). Cardiovascular dangers of air pollution. *J. Clin. Cardiol. Interv.* 2.

Ajit, A., Acharya, K., and Samanta, A. (2020). "A review of convolutional neural networks," in 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Germany, 2020, February, 1–5.

Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). "Understanding of a convolutional neural network," in 2017 International Conference on Engineering and Technology (ICET), USA, 2017, August, 1–6.

An, G., Xing, M., He, B., Liao, C., Huang, X., Shang, J., et al. (2020). Using machine learning for estimating rice chlorophyll content from *in situ* hyperspectral data. *Remote Sens.* 12 (18), 3104. doi:10.3390/rs12183104

Aw, J., and Kleeman, M. J. (2003). Evaluating the first-order effect of intraannual temperature variability on urban air pollution. *J. Geophys. Res. Atmos.* 108 (D12). doi:10.1029/2002jd002688

Bank, D., Koenigstein, N., and Giryes, R. (2020). *Autoencoders. arXiv preprint arXiv: 2003.05991*.

Bekkar, A., Hssina, B., Douzi, S., and Douzi, K. (2021). Air-pollution prediction in smart city, deep learning approach. *J. Big Data* 8 (1), 161–221. doi:10.1186/s40537-021-00548-1

Bell, L. N. (2020). Moisture effects on food's chemical stability. *Water Activity Foods Fundam. Appl.*, 227–253. doi:10.1002/9781118765982.ch9

Bell, M. L. (2006). The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia Region of the United States. *Environ. Int.* 32 (5), 586–593. doi:10.1016/j.envint.2006.01.005

Bergstra, A. D., Brunekreef, B., and Burdorf, A. (2018). The effect of industry-related air pollution on lung function and respiratory symptoms in school children. *Environ. Health* 17 (1), 30–39. doi:10.1186/s12940-018-0373-2

Bhargav, A. (2020). Air pollution-sources and classification. *Op Acc J Bio Sci Res* 1 (4). doi:10.46718/jbgsr.2020.01.000022

Briggs, D. J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., et al. (1997). Mapping urban air pollution using GIS: a regression-based approach. *Int. J. Geogr. Inf. Sci.* 11 (7), 699–718. doi:10.1080/136588197242158

Bui, T. C., Le, V. D., and Cha, S. K. (2018). A deep learning approach for forecasting air pollution in South Korea using LSTM. doi:10.48550/arXiv.1804.07891

Castelli, M., Clemente, F. M., Popovič, A., Silva, S., and Vanneschi, L. (2020). A machine learning approach to predict air quality in California. *Complexity* 2020, 1–23. doi:10.1155/2020/8049504

Chai, T., and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* 7 (3), 1247–1250. doi:10.5194/gmd-7-1247-2014

Chen, T. M., Kuschner, W. G., Gokhale, J., and Shofer, S. (2007). Outdoor air pollution: nitrogen dioxide, sulfur dioxide, and carbon monoxide health effects. *Am. J. Med. Sci.* 333 (4), 249–256. doi:10.1097/maj.0b013e31803b900f

Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., et al. (2018). GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci. Total Environ.* 634, 853–867. doi:10.1016/j.scitotenv.2018.04.055

Cheng, Z., Sun, H., Takeuchi, M., and Katto, J. (2018). "Deep convolutional autoencoder-based lossy image compression," in *2018 picture coding symposium (PCS)* (Germany: IEEE), 253–257.

Chicco, D., Warrens, M. J., and Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* 7, e623. doi:10.7717/peerj-cs.623

Dairi, A., Harrou, F., Khadraoui, S., and Sun, Y. (2021). Integrated multiple directed attention-based deep learning for improved air pollution forecasting. *IEEE Trans. Instrum. Meas.* 70, 1–15. doi:10.1109/tim.2021.3091511

Davis, J., and Goadrich, M. (2006). "The relationship between Precision-Recall and ROC curves," in Proceedings of the 23rd International Conference on Machine Learning, China, 2006, June, 233–240.

Delavar, M. R., Gholami, A., Shiran, G. R., Rashidi, Y., Nakhaeizadeh, G. R., Fedra, K., et al. (2019). A novel method for improving air pollution prediction based on machine learning approaches: a case study applied to the capital city of Tehran. *ISPRS Int. J. Geo-Information* 8 (2), 99. doi:10.3390/ijgi8020099

Dons, E., Van Poppel, M., Panis, L. I., De Prins, S., Berghmans, P., Koppen, G., et al. (2014). Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Sci. Total Environ.* 476, 378–386. doi:10.1016/j.scitotenv.2014.01.025

Dragomir, C. M., Voiculescu, M., Constantin, D. E., and Georgescu, L. P. (2015). Prediction of the NO2 concentration data in an urban area using multiple regression and neuronal networks AIP Conference Proceedings. *AIP Publ. LLC* 1694 (1), 040003. doi:10.1063/1.4937255

Endeshaw, M. F. L., and Endeshaw, L. (2020). Influence of temperature and relative humidity on air pollution in addis ababa, Ethiopia. *J. Environ. Earth Sci.* 2 (02), 19–25. doi:10.30564/jees.v2i2.2286

Farahani, M., Razavi-Termeh, S. V., and Sadeghi-Niaraki, A. (2022). A spatially based machine learning algorithm for potential mapping of the hearing senses in an urban environment. *Sustain. Cities Soc.* 80, 103675. doi:10.1016/j.scs.2022.103675

Faridi, S., Niazi, S., Yousefian, F., Azimi, F., Pasalari, H., Momeniha, F., et al. (2019). Spatial homogeneity and heterogeneity of ambient air pollutants in Tehran. *Sci. Total Environ.* 697, 134123. doi:10.1016/j.scitotenv.2019.134123

Garg, A., and Tai, K. (2013). Comparison of statistical and machine learning methods in modelling of data with multicollinearity. *Int. J. Model. Identif. Control* 18 (4), 295–312. doi:10.1504/ijmic.2013.053535

Ge, Y., Fu, Q., Yi, M., Chao, Y., Lei, X., Xu, X., et al. (2022). High spatial resolution land-use regression model for urban ultrafine particle exposure assessment in Shanghai, China. *Sci. Total Environ.* 816, 151633. doi:10.1016/j.scitotenv.2021.151633

Ghorbanzadeh, O., Meena, S. R., Blaschke, T., and Aryal, J. (2019). UAV-based slope failure detection using deep-learning convolutional neural networks. *Remote Sens.* 11 (17), 2046. doi:10.3390/rs11172046

Gorai, A. K., Tuluri, F., Tchounwou, P. B., and Ambinakudige, S. (2015). Influence of local meteorology and NO2 conditions on ground-level ozone concentrations in the eastern part of Texas, USA. *Air Qual. Atmos. Health* 8, 81–96. doi:10.1007/s11869-014-0276-5

Guarnieri, M., and Balmes, J. R. (2014). Outdoor air pollution and asthma. *Lancet* 383 (9928), 1581–1592. doi:10.1016/s0140-6736(14)60617-6

Gulliver, J., and Briggs, D. (2011). STEMS-Air: a simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci. Total Environ.* 409 (12), 2419–2429. doi:10.1016/j.scitotenv.2011.03.004

Guo, R., and Jiang, Y. (2020). Effects of precipitation on air pollution in spring and summer in lanzhou. *E3S Web Conf.* 194, 04007. doi:10.1051/e3sconf/202019404007

He, H., Wang, Y., Ma, Q., Ma, J., Chu, B., Ji, D., et al. (2014). Mineral dust and NOx promote the conversion of SO2 to sulfate in heavy pollution days. *Sci. Rep.* 4 (1), 4172. doi:10.1038/srep04172

Hogland, J., and Anderson, N. (2017). Function modeling improves the efficiency of spatial modeling using big data from remote sensing. *Big Data Cognitive Comput.* 1, 3. doi:10.3390/bdcc1010003

Hong, Q., Liu, C., Hu, Q., Xing, C., Tan, W., Liu, T., et al. (2021). Vertical distributions of tropospheric SO2 based on MAX-DOAS observations: investigating the impacts of regional transport at different heights in the boundary layer. *J. Environ. Sci.* 103, 119–134. doi:10.1016/j.jes.2020.09.036

Hosseini, V., and Shahbazi, H. (2016). Urban air pollution in Iran. *Iran. Stud.* 49 (6), 1029–1046. doi:10.1080/00210862.2016.1241587

Hu, K., Sivaraman, V., Bhrugubanda, H., Kang, S., and Rahman, A. (2016). "SVR based dense air pollution estimation model using static and wireless sensor network," in *2016 ieee sensors* (China: IEEE), 1–3.

Hu, X., Chu, L., Pei, J., Bian, J., and Liu, W. (2021). *Deep learning model complexity: concepts and approaches.*

Janhäll, S. (2015). Review on urban vegetation and particle air pollution – deposition and dispersion. *Atmos. Environ.* 105, 130–137. doi:10.1016/j.atmosenv.2015.01.052

Jia, W. J., Wang, M. F., Zhou, C. H., and Yang, Q. H. (2021). Analysis of the spatial association of geographical detector-based landslides and environmental factors in the southeastern Tibetan Plateau, China. *PLOS ONE* 16 (5), e0251776. doi:10.1371/journal.pone.0251776

Jiang, Z., Zheng, T., Bergin, M., and Carlson, D. (2022). Improving spatial variation of ground-level PM2.5 prediction with contrastive learning from satellite imagery. *Sci. Remote Sens.* 5, 100052. doi:10.1016/j.srs.2022.100052

Kalajdjieski, J., Zdravevski, E., Corizzo, R., Lameski, P., Kalajdziski, S., Pires, I. M., et al. (2020) Air pollution prediction with multi-modal data and deep neural networks. *Remote Sens.* 12, 4142, doi:10.3390/rs12244142

Kayes, I., Shahriar, S. A., Hasan, K., Akhter, M., Kabir, M. M., and Salam, M. A. (2019). The relationships between meteorological parameters and air pollutants in an urban environment. *Glob. J. Environ. Sci. Manag.* 5 (3), 265–278.doi:10.22034/GJESM.2019.03.01

Khosravi, K., Panahi, M., Golkarian, A., Keesstra, S. D., Saco, P. M., Bui, D. T., et al. (2020). Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran. *J. Hydrology* 591, 125552. doi:10.1016/j.jhydrol.2020.125552

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean J. Anesthesiol.* 72 (6), 558–569. doi:10.4097/kja.19087

Kumar, A., Gupta, I., Brandt, J., Kumar, R., Dikshit, A. K., and Patil, R. S. (2016). Air quality mapping using GIS and economic evaluation of health impact for Mumbai city, India. *J. Air & Waste Manag. Assoc.* 66 (5), 470–481. doi:10.1080/10962247.2016.1143887

Li, X., Ma, Y., Wang, Y., Wei, W., Zhang, Y., Liu, N., et al. (2019). Vertical distribution of particulate matter and its relationship with planetary boundary layer structure in Shenyang, Northeast China. *Aerosol Air Qual. Res.* 19 (11), 2464–2476. doi:10.4209/aaqr.2019.06.0311

Liu, Y., Zhou, Y., and Lu, J. (2020). Exploring the relationship between air pollution and meteorological conditions in China under environmental governance. *Sci. Rep.* 10 (1), 14518–14611. doi:10.1038/s41598-020-71338-7

Lü, G., Batty, M., Strobl, J., Lin, H., Zhu, A. X., and Chen, M. (2019). Reflections and speculations on the progress in Geographic Information Systems (GIS): a geographic perspective. *Int. J. Geogr. Inf. Sci.* 33 (2), 346–367. doi:10.1080/13658816.2018.1533136

Lu, H., Fu, X., Liu, C., Li, L. G., He, Y. X., and Li, N. W. (2017). Cultivated land information extraction in UAV imagery based on deep convolutional neural network and transfer learning. *J. Mt. Sci.* 14 (4), 731–741. doi:10.1007/s11629-016-3950-2

Lv, Y., Duan, Y., Kang, W., Li, Z., and Wang, F.-Y. (2014). Traffic flow prediction with big data: a deep learning approach. *IEEE Trans. Intelligent Transp. Syst.* 16, 865–873. doi:10.1109/TITS.2014.2345663

Ma, R., Liu, N., Xu, X., Wang, Y., Noh, H. Y., Zhang, P., et al. (2019). "A deep autoencoder model for pollution map recovery with mobile sensing networks," in Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, USA, 2019 July, 577–583.

Makarovskikh, T., and Herreinstein, A. (2022). *Urban air pollution modeling: a critical review*.

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., and Bezirtzoglou, E. (2020). Environmental and health impacts of air pollution: a review. *Front. public health* 8, 14. doi:10.3389/fpubh.2020.00014

Matejko, M., Dore, A. J., Hall, J., Dore, C. J., Błaś, M., Kryza, M., et al. (2009). The influence of long term trends in pollutant emissions on deposition of sulphur and nitrogen and exceedance of critical loads in the United Kingdom. *Environ. Sci. Policy* 12, 882–896. doi:10.1016/j.envsci.2009.08.005

Mavroidis, I., and Ilia, M. (2012). Trends of NOx, NO2 and O3 concentrations at three different types of air quality monitoring stations in Athens, Greece. *Atmos. Environ.* 63, 135–147. doi:10.1016/j.atmosenv.2012.09.030

Mavroulidou, M., Hughes, S. J., and Hellawell, E. E. (2004). A qualitative tool combining an interaction matrix and a GIS to map vulnerability to traffic induced air pollution. *J. Environ. Manag.* 70, 283–289. doi:10.1016/j.jenvman.2003.12.002

Mölter, A., Lindley, S., De Vocht, F., Simpson, A., and Agius, R. (2010). Modelling air pollution for epidemiologic research–Part II: predicting temporal variation through land use regression. *Sci. total Environ.* 409, 211–217. doi:10.1016/j.scitotenv.2010.10.005

Mombeini, H., and Yazdani-Chamzini, A. (2015). Modeling gold price via artificial neural network. *J. Econ. Bus. Manag.* 3, 699–703. doi:10.7763/joebm.2015.v3.269

Naddafi, K., Hassanvand, M. S., Yunesian, M., Momeniha, F., Nabizadeh, R., Faridi, S., et al. (2012). Health impact assessment of air pollution in megacity of Tehran, Iran. *Iran. J. Environ. health Sci. Eng.* 9, 28–37. doi:10.1186/1735-2746-9-28

Nowak, D. J., Hirabayashi, S., Bodine, A., and Hoehn, R. (2013). Modeled PM2.5 removal by trees in ten US cities and associated health effects. *Environ. Pollut.* 178, 395–402. doi:10.1016/j.envpol.2013.03.050

Noyes, P. D., McElwee, M. K., Miller, H. D., Clark, B. W., Van Tiem, L. A., Walcott, K. C., et al. (2009). The toxicology of climate change: environmental contaminants in a warming world. *Environ. Int.* 35, 971–986. doi:10.1016/j.envint.2009.02.006

Oji, S., and Adamu, H. (2020). Correlation between air pollutants concentration and meteorological factors on seasonal air quality variation. *J. air Pollut. health* 5, 11–32. doi:10.18502/japh.v5i1.2856

Oleniacz, R., Bogacki, M., Szulecka, A., Rzeszutek, M., and Mazur, M. (2016). Assessing the impact of wind speed and mixing-layer height on air quality in Krakow (Poland) in the years 2014–2015. *JCEEA* 33, 315–342. doi:10.7862/rb.2016.168

Özbay, B. (2012). Modeling the effects of meteorological factors on SO2 and PM10 concentrations with statistical approaches. *Clean–Soil, Air, Water* 40, 571–577. doi:10.1002/clen.201100356

Peng, Z.-R., Wang, D., Wang, Z., Gao, Y., and Lu, S. (2015). A study of vertical distribution patterns of PM2. 5 concentrations based on ambient monitoring with unmanned aerial vehicles: a case in Hangzhou, China. *Atmos. Environ.* 123, 357–369. doi:10.1016/j.atmosenv.2015.10.074

Pham, V. D., Nguyen, Q.-H., Nguyen, H.-D., Pham, V.-M., and Bui, Q.-T. (2020). Convolutional neural network—optimized moth flame algorithm for shallow landslide susceptible analysis. *IEEE Access* 8, 32727–32736. doi:10.1109/access.2020.2973415

Pinaya, W. H. L., Vieira, S., Garcia-Dias, R., and Mechelli, A. (2020). *Autoencoders, machine learning*. Germany: Elsevier, 193–208.

Quercia, D., Schifanella, R., Aiello, L. M., and McLean, K. (2015). Smelly maps: the digital life of urban smellscapes. *Proc. Int. AAAI Conf. Web Soc. Media* 9, 327–336. doi:10.1609/icwsm.v9i1.14621

Razavi-Termeh, S. V., Sadeghi-Niaraki, A., and Choi, S.-M. (2021). Spatial modeling of asthma-prone areas using remote sensing and ensemble machine learning algorithms. *Remote Sens.* 13, 3222. doi:10.3390/rs13163222

Richmond-Bryant, J., Chris Owen, R., Graham, S., Snyder, M., McDow, S., Oakes, M., et al. (2017). Estimation of on-road NO 2 concentrations, NO 2/NO X ratios, and related roadway gradients from near-road monitoring data. *Air Qual. Atmos. Health* 10, 611–625. doi:10.1007/s11869-016-0455-7

Richter, A., Burrows, J. P., Nüß, H., Granier, C., and Niemeier, U. (2005). Increase in tropospheric nitrogen dioxide over China observed from space. *Nature* 437, 129–132. doi:10.1038/nature04092

Rodrigue, J.-P. (2020). *The geography of transport systems*. USA: Routledge.

Sakti, A. D., Anggraini, T. S., Ihsan, K. T. N., Misra, P., Trang, N. T. Q., Pradhan, B., et al. (2023). Multi-air pollution risk assessment in Southeast Asia region using integrated remote sensing and socio-economic data products. *Sci. Total Environ.* 854, 158825. doi:10.1016/j.scitotenv.2022.158825

Salmond, J., Williams, D., Laing, G., Kingham, S., Dirks, K., Longley, I., et al. (2013). The influence of vegetation on the horizontal and vertical distribution of pollutants in a street canyon. *Sci. total Environ.* 443, 287–298. doi:10.1016/j.scitotenv.2012.10.101

Samet, J. M. (2007). Traffic, air pollution, and health. *Inhal. Toxicol.* 19, 1021–1027. doi:10.1080/08958370701533541

Santosa, S. J., Okuda, T., and Tanaka, S. (2008). Air pollution and urban air quality management in Indonesia. *CLEAN–Soil, Air, Water* 36, 466–475. doi:10.1002/clen.200800038

Sengupta, S., Patil, R., and Venkatachalam, P. (1996). Assessment of population exposure and risk zones due to air pollution using the geographical information system. *Comput. Environ. urban Syst.* 20, 191–199. doi:10.1016/s0198-9715(96)00014-2

Setälä, H., Viippola, V., Rantalainen, A.-L., Pennanen, A., and Yli-Pelkonen, V. (2013). Does urban vegetation mitigate air pollution in northern conditions? *Environ. Pollut.* 183, 104–112. doi:10.1016/j.envpol.2012.11.010

Sewani, H., and Kashef, R. (2020). An autoencoder-based deep learning classifier for efficient diagnosis of autism. *Children* 7, 182. doi:10.3390/children7100182

Shankar, V., and Parsana, S. (2022). An overview and empirical comparison of natural language processing (NLP) models and an introduction to and empirical application of autoencoder models in marketing. *J. Acad. Mark. Sci.* 50, 1324–1350. doi:10.1007/s11747-022-00840-3

Shi, Y., Bilal, M., Ho, H. C., and Omar, A. (2020). Urbanization and regional air pollution across South Asian developing countries–A nationwide land use regression for ambient PM2. 5 assessment in Pakistan. *Environ. Pollut.* 266, 115145. doi:10.1016/j.envpol.2020.115145

Shogrkhodaei, S. Z., Razavi-Termeh, S. V., and Fathnia, A. (2021). Spatio-temporal modeling of PM2. 5 risk mapping using three machine learning algorithms. *Environ. Pollut.* 289, 117859. doi:10.1016/j.envpol.2021.117859

Shukla, J., Misra, A., Sundar, S., and Naresh, R. (2008). Effect of rain on removal of a gaseous pollutant and two different particulate matters from the atmosphere of a city. *Math. Comput. Model.* 48, 832–844. doi:10.1016/j.mcm.2007.10.016

Somvanshi, S. S., Vashisht, A., Chandra, U., and Kaushik, G. (2019). Delhi air pollution modeling using remote sensing technique. *Handb. Environ. Mater. Manag.*, 1–27. doi:10.1007/978-3-319-58538-3_174-1

Steininger, M., Kobs, K., Zehe, A., Lautenschlager, F., Becker, M., and Hotho, A. (2020). Maplur: exploring a new paradigm for estimating air pollution using deep learning on map images. *ACM Trans. Spatial Algorithms Syst. (TSAS)* 6, 1–24. doi:10.1145/3380973

Su, H., Cheng, Y., and Poschl, U. (2020). New multiphase chemical processes influencing atmospheric aerosols, air quality, and climate in the anthropocene. *Accounts Chem. Res.* 53, 2034–2043. doi:10.1021/acs.accounts.0c00246

Swamy, Y., Venkanna, R., Nikhil, G., Chitanya, D., Sinha, P., Ramakrishna, M., et al. (2012). Impact of nitrogen oxides, volatile organic compounds and black carbon on

atmospheric ozone levels at a semi arid urban site in Hyderabad. *Aerosol Air Qual. Res.* 12, 662–671. doi:10.4209/aaqr.2012.01.0019

Szopińska, K., Cienciała, A., Bieda, A., Kwiecień, J., Kulesza, Ł., and Parzych, P. (2022). Verification of the perception of the local community concerning air quality using ADMS-roads modeling. *Int. J. Environ. Res. Public Health* 19, 10908. doi:10.3390/ijerph191710908

Turalıoğlu, F. S., Nuhoğlu, A., and Bayraktar, H. (2005). Impacts of some meteorological parameters on SO2 and TSP concentrations in Erzurum, Turkey. *Chemosphere* 59, 1633–1642. doi:10.1016/j.chemosphere.2005.02.003

U.S. EPA (1978). *Altitude as a factor in air pollution*. Washington, D.C.: U.S. Environmental Protection Agency. EPA/600/9-78/015 (NTIS PB285645).

Vakkilainen, E. K. (2017). Solid biofuels and combustion. *Steam generation biomass*, 18–56. doi:10.1016/b978-0-12-804389-9.00002-2

Wang, F., Chen, D., Cheng, S., Li, J., Li, M., and Ren, Z. (2010). Identification of regional atmospheric PM10 transport pathways using HYSPLIT, MM5-CMAQ and synoptic pressure pattern analysis. *Environ. Model. Softw.* 25, 927–934. doi:10.1016/j.envsoft.2010.02.004

Wang, J., Tang, J., Xu, Z., Wang, Y., Xue, G., Zhang, X., et al. (2017). "Spatiotemporal modeling and prediction in cellular networks: a big data enabled deep learning approach," in *IEEE INFOCOM 2017-IEEE conference on computer communications* (USA: IEEE), 1–9.

Wang, J., and Xu, C. (2017). Geodetector: principle and prospective. *Acta Geogr. Sin.* 72, 116–134. doi:10.11821/dlxb201701010

Xu, H., Bechle, M. J., Wang, M., Szpiro, A. A., Vedal, S., Bai, Y., et al. (2019). National PM2. 5 and NO2 exposure models for China based on land use regression, satellite measurements, and universal kriging. *Sci. Total Environ.* 655, 423–433. doi:10.1016/j.scitotenv.2018.11.125

Xu, X., Qin, N., Zhao, W., Tian, Q., Si, Q., Wu, W., et al. (2022). A three-dimensional LUR framework for PM2. 5 exposure assessment based on mobile unmanned aerial vehicle monitoring. *Environ. Pollut.* 301, 118997. doi:10.1016/j.envpol.2022.118997

Yerramilli, A., Dodla, V. B. R., and Yerramilli, S. (2011). Air pollution, modeling and GIS based decision support systems for air quality risk assessment. *Adv. air Pollut.*, 295–324. doi:10.5772/22055

Yousefian, F., Faridi, S., Azimi, F., Aghaei, M., Shamsipour, M., Yaghmaeian, K., et al. (2020). Temporal variations of ambient air pollutants and meteorological influences on their concentrations in Tehran during 2012–2017. *Sci. Rep.* 10, 292. doi:10.1038/s41598-019-56578-6

Zaini, Na, Ean, L. W., Ahmed, A. N., Abdul Malek, M., and Chow, M. F. (2022). PM2. 5 forecasting for an urban area based on deep learning and decomposition method. *Sci. Rep.* 12, 17565. doi:10.1038/s41598-022-21769-1

Zavrak, S., and Iskefiyeli, M. (2020). Anomaly-based intrusion detection from network flow features using variational autoencoder. *IEEE Access* 8, 108346–108358. doi:10.1109/access.2020.3001350

Zhang, K., and Batterman, S. (2013). Air pollution and health risks due to vehicle traffic. *Sci. total Environ.* 450, 307–316. doi:10.1016/j.scitotenv.2013.01.074

Zhang, S., Zhou, Y., Yu, Y., Li, F., Zhang, R., and Li, W. (2022). Using the geodetector method to characterize the spatiotemporal dynamics of vegetation and its interaction with environmental factors in the qinba mountains, China. *Remote Sens.* 14, 5794. doi:10.3390/rs14225794

Zhao, K., Bao, Y., Huang, J., Wu, Y., Moshary, F., Arend, M., et al. (2019). A high-resolution modeling study of a heat wave-driven ozone exceedance event in New York City and surrounding regions. *Atmos. Environ.* 199, 368–379. doi:10.1016/j.atmosenv.2018.10.059

Zou, B., Peng, F., Wan, N., Mamady, K., and Wilson, G. J. (2014). Spatial cluster detection of air pollution exposure inequities across the United States. *PLoS One* 9, e91917. doi:10.1371/journal.pone.0091917

frontiers | Frontiers in Built Environment

TYPE Systematic Review
PUBLISHED 30 July 2024
DOI 10.3389/fbuil.2024.1321634

OPEN ACCESS

EDITED BY
Sayali Sandbhor,
Symbiosis International University, India

REVIEWED BY
Pritesh Shah,
Symbiosis International University, India
Sanjay Kulkarni,
Symbiosis International University, India

*CORRESPONDENCE
Reda Alhajj,
alhajj@ucalgary.ca

RECEIVED 16 October 2023
ACCEPTED 26 June 2024
PUBLISHED 30 July 2024

CITATION
Kaveh H and Alhajj R (2024), Recent advances in crack detection technologies for structures: a survey of 2022-2023 literature.
Front. Built Environ. 10:1321634.
doi: 10.3389/fbuil.2024.1321634

Recent advances in crack detection technologies for structures: a survey of 2022-2023 literature

Hessam Kaveh[1] and Reda Alhajj[1,2,3]*

[1]Department of Computer Engineering, Istanbul Medipol University, Istanbul, Türkiye, [2]Department of Computer Science, University of Calgary, Calgary, AB, Canada, [3]Department of Health Informatics, University of Southern Denmark, Odense, Denmark

Introduction: Cracks, as structural defects or fractures in materials like concrete, asphalt, and metal, pose significant challenges to the stability and safety of various structures. Addressing crack detection is of paramount importance due to its implications for public safety, infrastructure integrity, maintenance costs, asset longevity, preventive maintenance, economic impact, and environmental considerations.

Methods: In this survey paper, we present a comprehensive analysis of recent advancements and developments in crack detection technologies for structures, with a specific focus on articles published between 2022 and 2023. Our methodology involves an exhaustive search of the Scopus database using keywords related to crack detection and machine learning techniques. Among the 129 papers reviewed, 85 were closely aligned with our research focus.

Results: We explore datasets that underpin crack detection research, categorizing them as public datasets, papers with their own datasets, and those using a hybrid approach. The prevalence and usage patterns of public datasets are presented, highlighting datasets like Crack500, Crack Forest Dataset (CFD), and Deep Crack. Furthermore, papers employing proprietary datasets and those combining public and proprietary sources are examined. The survey comprehensively investigates the algorithms and methods utilized, encompassing CNN, YOLO, UNet, ResNet, and others, elucidating their contributions to crack detection. Evaluation metrics such as accuracy, precision, recall, F1-score, and IoU are discussed in the context of assessing model performance. The results of the 85 papers are summarized, demonstrating advancements in crack detection accuracy, efficiency, and applicability.

Discussion: Notably, we observe a trend towards using modern and novel algorithms, such as Vision Transformers (ViT), and a shift away from traditional methods. The conclusion encapsulates the current state of crack detection research, highlighting the integration of multiple algorithms, expert models, and innovative data collection techniques. As a future direction, the adoption of emerging algorithms like ViT is suggested. This survey paper serves as a

valuable resource for researchers, practitioners, and engineers working in the field of crack detection, offering insights into the latest trends, methodologies, and challenges.

KEYWORDS

Cracks, structural defects, infrastructure integrity, preventive maintenance, economic impact, survey

1 Introduction

Cracks are structural defects or fractures that occur in various materials, such as concrete, asphalt, and metal, often caused by stress, environmental factors, or wear over time. These imperfections can significantly compromise the integrity of structures, such as buildings, bridges, roads, and other infrastructures (Jiya et al., 2016). Understanding and addressing cracks is of paramount importance due to the following reasons:

Safety Concerns: Cracks in structures can pose severe safety hazards to the public. They weaken the structural stability, making buildings and bridges susceptible to collapse, potentially leading to injuries or loss of life.

Infrastructure Integrity: The presence of cracks can undermine the overall structural integrity of essential infrastructure. As cracks propagate and grow, they can weaken load-bearing elements, causing irreversible damage and costly repairs if not addressed promptly.

Maintenance Costs: Unchecked cracks can escalate maintenance costs significantly. Small cracks, when detected early, are easier and cheaper to repair than allowing them to worsen and cause extensive damage, requiring more extensive and costly rehabilitation.

Asset Longevity: Effective crack detection and timely repairs can extend the lifespan of structures. By addressing cracks early on, the overall durability and longevity of buildings and infrastructure can be significantly improved.

Preventive Maintenance: Crack detection plays a crucial role in implementing preventive maintenance strategies. Early identification allows for targeted repairs or reinforcement, preventing the cracks from spreading and mitigating potential risks.

Economic Impact: Infrastructure failure due to undetected cracks can result in significant economic losses. Repairs and structural rehabilitation can be costly, and in severe cases, infrastructure failures can disrupt transportation, utilities, and daily activities, impacting productivity and economic stability.

Environmental Impact: Cracked structures may allow for water ingress, leading to corrosion of reinforcement and other components. Water infiltration can further exacerbate cracks and compromise the structural integrity, impacting the environment and potentially leading to water-related issues like mold growth.

Given these critical implications, crack detection assumes immense significance in maintaining public safety, preserving infrastructure assets, and ensuring the efficient and sustainable operation of modern societies. Timely and accurate crack detection methods are vital tools for engineers, researchers, and practitioners, helping them assess structural health and make informed decisions to enhance the safety and longevity of our built environment.

In this survey paper, we focus on recent advancements and new developments in crack detection technologies for structures, with a specific emphasis on articles published in the years 2022 and 2023. To compile our findings, we conducted a thorough search of the Scopus database using the keywords "crack detection," "building," "road," "pavement," and "concrete." The search was further refined to include articles related to machine learning and deep learning techniques. The language criterion was set to English to ensure the coherence and consistency of the gathered information.

Our search yielded a total of 129 papers, of which 85 were closely aligned with our research focus. From Table 1 we can see the complete detail about the information from Scopus.

These articles serve as the foundation for our survey, enabling us to analyze the state-of-the-art developments and trends in crack detection within the context of structures. Additionally, we employed VOSviewer, a specialized software tool for bibliometric analysis, to generate keyword cloud maps, providing a visual representation of the prominent terms and concepts in the selected articles which it is obvious in Figure 1. It is clear that "Deep Learning" and,"Convolutional Neural Networks (CNN)" have the most relation with our main search topic "Crack Detection". Moreover, we have extracted valuable insights from the charts depicting "Documents by year," "Documents per year by source," "Documents by search area," and "Documents by country or territory," which contribute to our comprehensive understanding of the current landscape of crack detection research. As we can see from Figure 2 there are 78 papers which published in 2022 and 51 papers in 2023 up to now. We can see increment in the amount of the papers which published in "Sensors" and "Remote Sensing" in 2023 compare to 2022 due to the Figure 3. It is obviously clear that this search topic appears most in the fields of "Engineering" and "Computer Science", we can see this point from Figure 4. And from Figure 5 we can see the First 10 countries which published the papers in this field, more than the others.

Continuing with the survey paper, the subsequent sections delve into the methodologies used for crack detection. We explore traditional image processing techniques, such as edge detection, thresholding, and binary image analysis, highlighting their strengths and limitations. Additionally, we delve into the application of state-of-the-art deep learning architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Transformers, and other deep learning models, showcasing their superior performance and ability to capture intricate crack patterns (Golding et al., 2022; BaniMustafa et al., 2023).

Moving forward, we discuss potential future research directions in the field of crack detection. This includes the necessity for more diverse and comprehensive datasets, encompassing various types of structures, lighting conditions, and crack patterns (Sun et al., 2023).

TABLE 1 Summary of scopus query and search results.

Query in Scopus	(TITLE-ABS-KEY (“crack detection” “deep learning”) OR TITLE-ABS-KEY (“crack detection” “MachineLearning”) OR TITLE-ABS-KEY (“crack detection” “building” “road” “pavement” “concrete”)) AND PUBYEAR >2021 AND (LIMIT-TO (OA, “all”)) AND (LIMIT-TO (PUBSTAGE, “final”)) AND (LIMIT-TO (DOCTYPE, “ar”)) AND (LIMIT-TO (LANGUAGE, “English”))
Search Results	129
Unrelated Results	44
Related Results	85

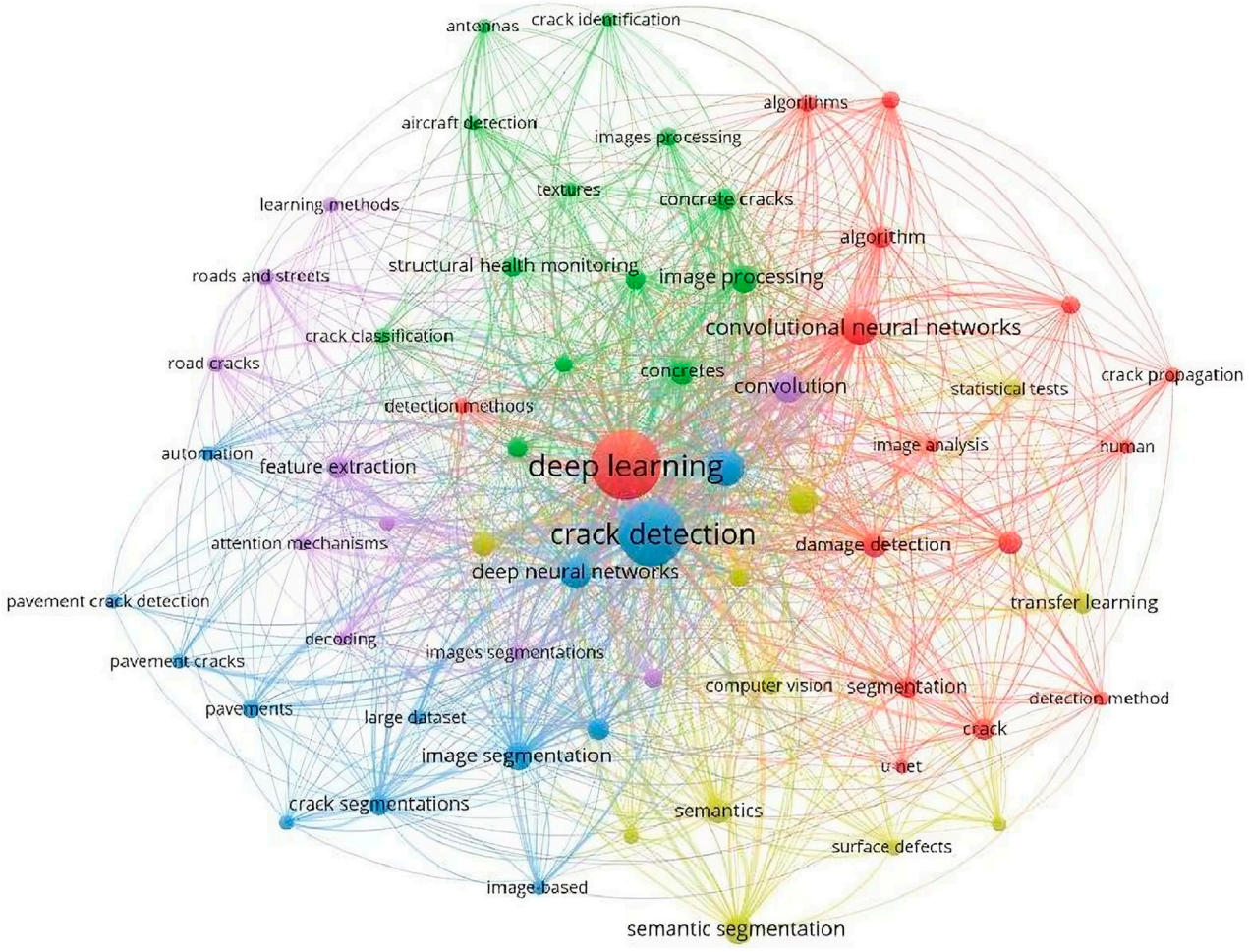


FIGURE 1
Keywords map by VOSviewer.

The survey paper also advocates for the development of real-time crack detection systems and the incorporation of explainable AI techniques to enhance the interpretability of crack detection models (Li G. et al., 2022; Ma D. et al., 2022).

In conclusion, this survey paper aims to be a valuable resource, consolidating the current knowledge on crack detection in structures. By reviewing both conventional and advanced techniques and providing insights into potential future developments, we aspire to inspire further advancements in this critical area, ultimately contributing to the safety, reliability, and longevity of vital infrastructures. To facilitate comprehension throughout the paper, we provide a list of acronyms along with their expanded forms which you can find it in Table 2.

2 The approach

2.1 Datasets in crack detection

In the realm of crack detection for various structures, the availability of diverse and appropriately curated datasets holds a pivotal role in advancing research and innovation. These datasets serve as the

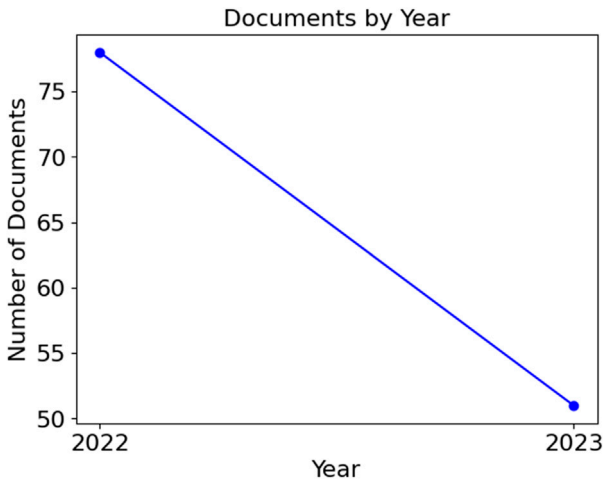


FIGURE 2
Documents per year by year.

foundation upon which crack detection models are trained, tested, and validated. A robust dataset ensures the realism and accuracy of the models, facilitating the development of effective and reliable crack detection methodologies. The creation and utilization of datasets in crack detection research encompass a wide spectrum of applications, from building and road infrastructure to pavements and concrete structures. These datasets encapsulate a range of crack types, sizes, orientations, and severity levels, mirroring the real-world scenarios that researchers and engineers encounter in practice. By incorporating the inherent complexity and variability of cracks, these datasets enable the evaluation and comparison of detection algorithms under diverse conditions. In the context of recent advancements, the integration of machine learning and deep learning techniques has spurred the demand for datasets that accommodate the unique challenges posed by these methods. Such datasets should not only represent structural defects accurately but also encompass a variety of environmental conditions, lighting variations, and perspectives, enhancing the models' adaptability and generalization capabilities. As we delve into the datasets utilized within the 85 papers reviewed in this survey, it becomes evident that researchers draw from a combination of sources to bolster the quality and comprehensiveness of their studies. Some papers leverage publicly available datasets, while others design and assemble their datasets, tailored to the specific objectives of their crack detection research. Moreover, a subset of studies combines both approaches, harnessing the power of existing public datasets and augmenting them with proprietary data to enhance the richness and diversity of training and testing scenarios. These datasets collectively contribute to the diversity and comprehensiveness of crack detection research. The subsequent sections of this paper delve into a detailed analysis of the datasets' origins and utilization. We categorize these datasets based on their sources and types, presenting a comprehensive overview of the dataset landscape. This analysis provides insights into the dataset selection and utilization practices that influence the evolution of crack detection technologies.

2.1.1 Public datasets

Publicly available datasets have played a crucial role in shaping the landscape of crack detection research. These datasets, carefully curated and made accessible to the research community, serve as valuable resources for benchmarking, testing, and validating crack detection algorithms. Researchers leverage these datasets to assess the performance of their methods and foster a collaborative environment

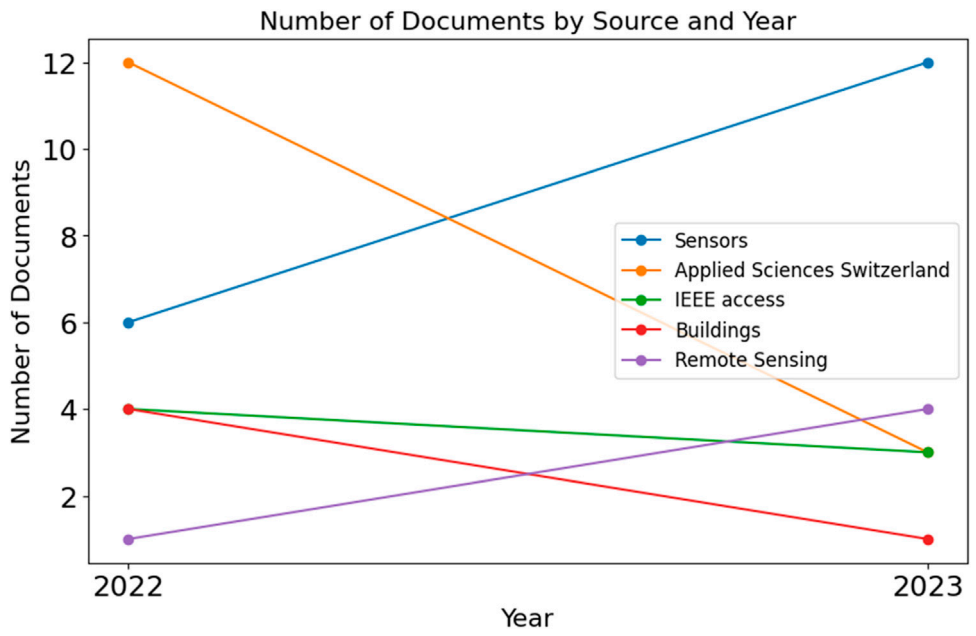


FIGURE 3
Documents per year by source.

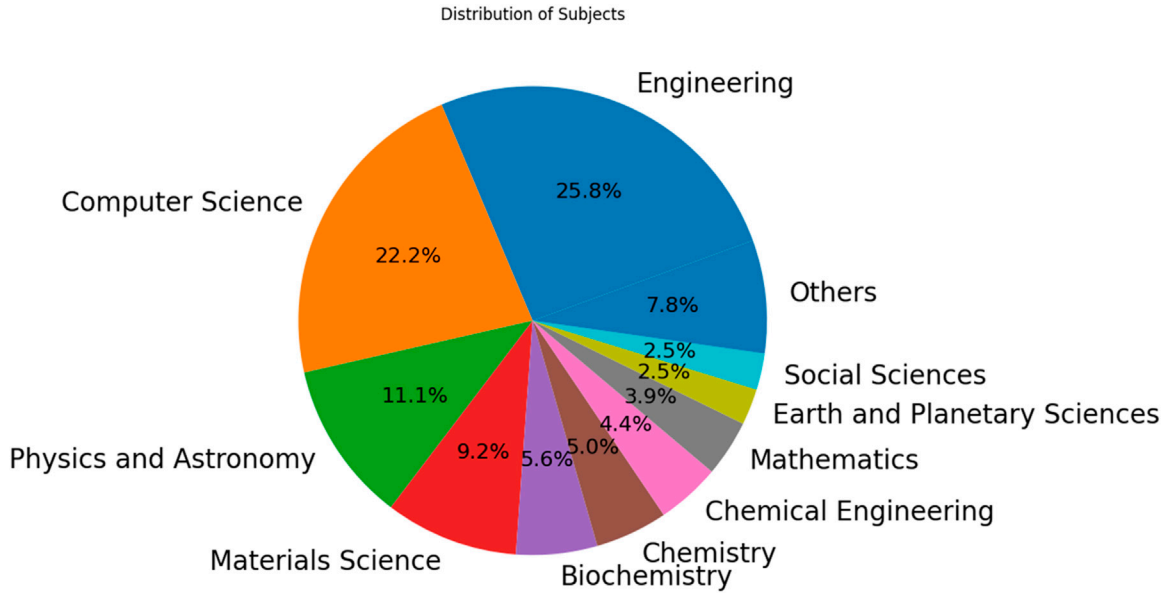


FIGURE 4
Documents by subject area.

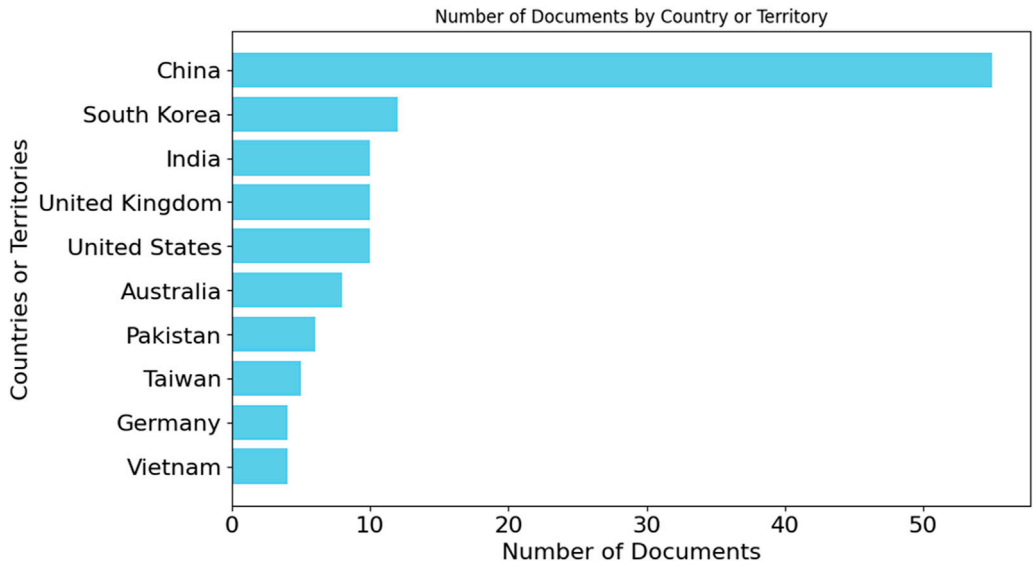


FIGURE 5
Documents by country or territory.

for advancing the field. In the pursuit of effective crack detection solutions, these datasets offer a diverse array of crack types, surface textures, and lighting conditions. The utilization of public datasets ensures a level playing field for researchers, enabling fair comparisons and promoting the development of innovative techniques. The following tables provide an overview of the distribution of papers per datasets and datasets per papers, shedding light on the prevalence and usage patterns of these publicly available resources. The datasets employed in these studies represent a diverse collection, each contributing to the advancement of crack detection technologies. Some of the prominent public datasets used in these papers include.

- Crack500: A dataset containing images of cracked and non-cracked concrete surfaces.
- Crack Forest Dataset (CFD): Images of cracked tree bark textures.
- CrackTree200: Images of tree bark with cracks for assessing detection techniques.
- Deep Crack: Images of various crack types for evaluating detection methods.
- GAPs384: Grayscale images containing cracks, patches, and non-defective areas for pavement crack detection.
- AigleRN: Images of cracks in road pavements for evaluation.
- CrackTree260: Tree bark images with cracks.

TABLE 2 List of acronyms used in this paper.

Acronym	Expanded form	Acronym	Expanded form
CNN	Convolutional Neural Network	ANN	Artificial Neural Network
FCN	Fully Convolutional Network	ViT	Vision Transformer
SSD	Single Shot MultiBox Detector	CRF	Conditional Random Field
IoU	Intersection over Union	Rsef	Regional-Scale Edge Feature
mAP	Mean Average Precision	ResNet	Residual Neural Network
ROC	Receiver Operating Characteristic	UNet	U-Net
PR	Precision-Recall	YOLO	You Only Look Once
MCC	Matthews Correlation Coefficient	kNN	k-Nearest Neighbors
DCNN	Deep Convolutional Neural Network	GAN	Generative Adversarial Network

- CRKWH100: Grayscale images of road surface cracks.
- CrackLS315: Images of cracked and non-cracked surfaces.
- RDD2022: Dataset focused on road damage detection, including cracks.
- DeepCrack537: Extended version of the Deep Crack dataset with a larger set of images.
- AED: Images of asphalt surfaces, including cracks.
- CrackSegNet: Dataset for evaluating the CrackSegNet model.
- DAGM 2007: Dataset used in image analysis and pattern recognition.
- CCIC: Dataset containing images of building cracks.
- CrackTree206: Additional tree bark images with cracks.
- SYCrack: Dataset used for crack detection research.
- Mixed Crack Dataset (MCD): Dataset containing mixed crack types.
- Building Wall Crack Images (BWCI): Images of cracks in building walls.
- SDNET2018: Comprehensive dataset for concrete crack detection.
- Crack45K: Large dataset with images of various crack types.
- Stone331: Dataset with images of stone surfaces and cracks.
- CQU-BPDD: Images of bridge pavement cracks.
- Historical Building Crack 2019: Dataset focusing on historical building cracks.

Table 3 shows the complete information about the public datasets which used in the studies.

2.1.2 Papers with their Own dataset

Some of the papers among the 85 reviewed in this survey have taken a proactive approach by creating their own datasets tailored to their research objectives. These researchers recognized the importance of aligning the dataset with the specific characteristics of their crack detection problem. By meticulously designing and curating their datasets, these studies aimed to capture the nuances of real-world scenarios, considering factors such as structural types, crack severity levels, lighting conditions, and surface textures. Creating a custom dataset offers several advantages. Researchers have the flexibility to control and manipulate variables to simulate a wide range of scenarios, contributing to a more controlled experimentation environment. Furthermore, custom datasets can address specific challenges or limitations present in publicly available datasets. However, this approach requires substantial effort in data collection, annotation, and validation, ensuring the dataset's integrity and applicability. The complete details of these proprietary datasets are presented in Table 4.

2.1.3 Papers using both their Own and public datasets

In the landscape of crack detection research, a subset of the reviewed papers stands out by leveraging a dual-source approach. These studies draw from the strengths of both their own meticulously curated datasets and publicly available datasets. By merging these sources, researchers aim to achieve a harmonious balance between dataset richness and diversity. The integration of proprietary and public datasets provides a unique opportunity for robust training and evaluation. Researchers benefit from the specificity and customization of their dataset while also capitalizing on the broader scope and variety offered by public datasets. This combination empowers researchers to validate the adaptability and generalization capabilities of their crack detection models across a spectrum of scenarios. Papers adopting this hybrid approach acknowledge the complementary nature of different datasets and recognize that collaborative efforts between proprietary and public sources can foster innovation and drive the advancement of crack detection techniques. For further insights into the specific datasets employed in these papers, refer to Table 5. These variations underscore the dynamic nature of crack detection research and highlight the multifaceted strategies researchers employ to overcome challenges and contribute to the evolution of structural health assessment technologies.

2.2 Algorithms and methods in Crack detection

The development and application of algorithms and methods in crack detection are central to the advancement of structural health assessment. As the field of crack detection has evolved, machine learning and deep learning techniques have emerged as powerful tools for automated crack detection, offering innovative solutions to the challenges posed by crack identification and characterization.

TABLE 3 Datasets and their brief explanation per papers.

Dataset	Explanation	Paper	Type	Download link
Crack500	Crack500 was curated by capturing 500 RGB color images featuring cracks on the surfaces of 500 asphalt roads, each with a resolution of 2560 × 1440. (Lee et al., 2023)	Munawar et al. (2022a), Mohammed et al. (2022), Yong and Wang (2022), Zhang et al. (2022), Zhao et al. (2023a), Lee et al. (2023), Yang et al. (2023)	Image	https://www.kaggle.com/datasets/pauldavid22/crack500202205091090436z001
Crack Forest Dataset (CFD)	The Crack Forest dataset comprises 118 pairs of RGB color images capturing asphalt road surface cracks in Beijing, China. These images were taken using an iPhone 5, and they maintain their original resolution of 480 × 320 pixels	Lee et al. (2023), Munawar et al. (2022b), Gooda et al. (2023), Inácio et al. (2023), Zhang et al. (2023a), Zhao et al. (2022), Jing et al. (2022), Zhao et al. (2023a), Li et al. (2022a)	Image	https://github.com/cuilimeng/CrackForest-dataset/tree/master
CrackTree200	CrackTree200 offers high-resolution images at 800 × 600 pixels along with corresponding label values identifying surface cracks on asphalt surfaces. The dataset includes numerous images featuring cracks on asphalt surfaces with tree shadows. (Lee et al., 2023)	Zhao et al. (2023a), Lee et al. (2023), Lv et al. (2023)	Image	https://github.com/fyangneil/pavement-crack-detection/tree/master?tab=readme-ov-file
Deep Crack	The crack segmentation dataset encompasses 537 RGB color images, each with dimensions of 554 × 384 pixels. This dataset is characterized by its inclusion of images at various scales, showcasing cracks that can occur on surfaces composed of different materials. It features images depicting cracks on both concrete and asphalt surfaces. (Lee et al., 2023)	Lee et al. (2023), Zhang et al. (2022), Panta et al. (2023), de León et al. (2023), Jing et al. (2022)	Image	https://github.com/yhlleo/DeepCrack
GAPs384	The German Asphalt Pavement Distress dataset comprises a collection of road surface images accompanied by labeling data, encompassing various distress types such as cracks, potholes, and inlaid paths. These images have a resolution of 1920 × 1,080 pixels, and the dataset consists of a total of 509 images. (Lee et al., 2023)	Li et al. (2022a), Munawar et al. (2022a), Zhao et al. (2023a), Lee et al. (2023), Lv et al. (2023)	Image	https://github.com/fyangneil/pavement-crack-detection/tree/master?tab=readme-ov-file
AigleRN	The dataset includes 38 pre-processed grayscale images depicting a road pavement surface located in France. This dataset is divided into two sets, with half of the images having dimensions of 991 × 462 pixels, and the remaining half having dimensions of 311 × 462 pixels. (Inácio et al., 2023)	Inácio et al. (2023)	Image	https://github.com/Sutadasuto/uvgg19_crack_detection?tab=readme-ov-file
CrackTree260	The dataset comprises 260 images, each with dimensions of 800 × 600 pixels, captured using an area-array camera under visible light illumination conditions. (Inácio et al., 2023)	Inácio et al. (2023), Siriborvornratanakul (2022), Wang et al. (2022a)	Image	https://github.com/qinnzou/DeepCrack
CRKWH100	The dataset includes 100 images of a road pavement surface, each with dimensions of 512 × 512 pixels, captured using a line array camera under visible light illumination conditions. (Inácio et al., 2023)	Inácio et al. (2023), Zhao et al. (2023b), Siriborvornratanakul (2022), Wang et al. (2022a)	Image	https://github.com/qinnzou/DeepCrack
CrackLS315	The dataset comprises 315 images, each with dimensions of 512 × 512 pixels, captured using a line array camera under laser illumination. (Inácio et al., 2023)	Inácio et al. (2023), Zhao et al. (2023b), Siriborvornratanakul (2022), Wang et al. (2022a)	Image	https://github.com/qinnzou/DeepCrack
RDD2022	The dataset includes 21,041 road damage images from Japan, the Czech Republic, and India, distributed as follows: 10,506 from Japanese pavement, 7,706 from Indian pavement, and 2,829 from Czech pavement. It covers various road conditions, comprising eight distinct categories, such as longitudinal cracks, transverse cracks, and crosswalk blur. (Yu and Zhou, 2023)	Liu et al. (2022), Ashraf et al. (2023), Yu and Zhou (2023)	Image	https://figshare.com/articles/dataset/RDD2022_-_The_multi-national_Road_Damage_Dataset_released_through_CRDDC_2022/21431547

(Continued on following page)

TABLE 3 (*Continued*) Datasets and their brief explanation per papers.

Dataset	Explanation	Paper	Type	Download link
DeepCrack537	This dataset consists of 537 images, each accompanied by annotated labels. All images and their corresponding labels share a uniform size of 544 × 384 pixels. (Zhang et al., 2023a)	Zhang et al. (2023a)	Image	Not Available
AED	The AED dataset consists of three sub-datasets: AigleRN (38 images), ESAR (15 images), and Dynamique (16 images). These images are unevenly illuminated and prone to noise interference, with mostly small cracks. (Zhang et al., 2023a)	Zhang et al. (2023a)	Image	https://universe.roboflow.com/ben-ohmju/aed-0awut/dataset/1
CrackSegNet	The CrackSegNet dataset contains 919 crack images. (Yang et al., 2023)	Yang et al. (2023)	Image	Not Available
DAGM 2007	The DAGM 2007 dataset is organized into 10 classes, with 6 classes designated for development and 4 classes for the 2007 DAGM sym- posium competition. In the initial 6 classes, there are 1,000 nonde-fective images and 150 defective images each, while the remaining 4 classes comprise 2,000 nondefective and 300 defective images per class. Each class is generated using distinct texture and defect models. (Kim et al., 2023b)	Kim et al. (2023b)	Image	https://www.kaggle.com/datasets/mhskjelvareid/dagm-2007-competition-dataset-optical-inspection
CCIC	This dataset comprises concrete images with cracks collected from various METU Campus Buildings, divided into negative and positive crack images for classification. Each class contains 20,000 images, totaling 40,000 images with dimensions of 227 × 227 pixels in RGB. It originates from 458 high-resolution images (4032 × 3024 pixels) using (Zhang et al., 2016)'s method. The high-resolution images exhibit variations in surface finish and illumination conditions. No data augmentation techniques, such as random rotation or flipping, are employed	Ozgenel and Sorguc¸ (2018), Wang et al. (2022b), Golding et al. (2022), Islam et al. (2022), Jayaraju et al. (2022), Paramanandham et al. (2022), Pu et al. (2022), Bai et al. (2023), Shim et al. (2023)	Image	https://www.kaggle.com/datasets/arnavr10880/concrete-crack-images-for-classification
CrackTree206	The dataset consists of 206 images of road cracks, each with a resolution of 800 × 600 pixels, and may contain occlusions and shadows. (Zhang et al., 2022)	Zhang et al. (2022)	Image	https://github.com/qinnzou/DeepCrack
SYCrack	There are 177 images, all sized uniformly at 256,256 pixels, which were captured from cracks on the Suoyang Ancient City wall surface. (Zhang et al., 2022)	Zhang et al. (2022)	Image	Not Available
Mixed Crack Dataset (MCD)	The MCD dataset comprises a total of 2,538 raw images along with their corresponding annotations. These images encompass concrete cracks, asphalt cracks, and earthen cracks, sourced from datasets such as CrackTree206, Crack500, Deep Crack, and SYCrack. (Zhang et al., 2022)	Zhang et al. (2022)	Image	Not Available
Building Wall Crack Images (BWCI)	BWCI consists of 4,500 wall crack images with 27 27 pixels. (Islam et al., 2022)	Islam et al. (2022)	Image	Not Available
SDNET2018	SDNET2018 is an extensive dataset consisting of more than 56,000 images capturing both cracked and non-cracked concrete bridge decks, walls, and pavements. Notably, the dataset encompasses cracks ranging from as narrow as 0.06 mm to as wide as 25 mm. (Dorafshan et al., 2018)	Lv et al. (2023), Ngo et al. (2023), Philip et al. (2023), Qayyum et al. (2023), Lu et al. (2022), Inam et al. (2023), Kao et al. (2023), Shim et al. (2023), Li et al. (2023a), Popli et al. (2023)	Image	https://www.kaggle.com/datasets/aniruddhsharma/structural-defects-network-concrete-crack-images

(Continued on following page)

TABLE 3 (*Continued*) Datasets and their brief explanation per papers.

Dataset	Explanation	Paper	Type	Download link
Crack45K	The dataset comprises 45,000 images, each with a resolution of 224 × 224 pixels, showcasing diverse pavement surfaces, both with and without cracks. (Ali et al., 2022)	Ali et al. (2022)	Image	Not Available
Stone331	The dataset consists of 331 grayscale images of stone surfaces, each with a size of 512,512 pixels. It's worth noting that the entire image may not contain the stone, so masks are provided to exclude predictions outside the stone area. (Konig et al., 2021; Siriborvornratanakul, 2022)	Siriborvornratanakul (2022)	Image	https://github.com/qinnzou/DeepCrack
CQU-BPDD	The CQU-BPDD consists of 60,056 bituminous pavement images captured by in-vehicle cameras on a professional pavement inspec- tion vehicle in southern China. Each image corresponds to a 2 × 3 meters pavement patch on highways and has a resolution of 1200900 pixels. The dataset includes seven distress types: transverse crack, massive crack, alligator crack, crack pouring, longitudinal crack, ravelling, repair, and normal pavement conditions. (Tang et al., 2021)	Liu et al. (2022)	Image	https://github.com/DearCaat/CQU-BPDD
Historical Building Crack 2019	This dataset comprises 3,886 images, including annotated RGB images, with 757 depicting cracks and 3,139 depicting non-crack conditions. The raw images were captured using a Canon camera (Canon EOS REBEL T3i) with a resolution of 5184 × 3456 pixels. These images feature historical buildings, such as the Mosque (Mas- jed) of Amir al-Maridani, situated in Sekat Al Werdani, El-Darb El-Ahmar, in the Cairo Governorate. (Yadav et al., 2022)	Yadav et al. (2022)	Image	https://data.mendeley.com/datasets/xfk99kpmj9/1

2.2.1 Machine learning methods

Machine learning techniques encompass a range of methodologies that enable computers to learn patterns and make predictions from data without being explicitly programmed. These methods leverage statistical algorithms to recognize patterns and trends, making them well-suited for analyzing crack patterns and textures in images. One key advantage of machine learning is its versatility in handling various types of data and extracting relevant features for crack detection. However, the effectiveness of traditional machine learning methods can be limited by their dependence on hand-crafted features and their inability to capture complex spatial relationships within cracks.

2.2.2 Deep learning

Deep learning, a subset of machine learning, has gained immense popularity in recent years due to its ability to automatically learn hierarchical representations from raw data. Deep learning models, such as convolutional neural networks (CNNs), have demonstrated remarkable success in image classification, object detection, and segmentation tasks. CNNs excel at capturing intricate features and patterns within images, making them particularly well-suited for crack detection. Deep learning methods often outperform traditional machine learning approaches by automatically learning relevant features from data, eliminating the need for manual feature engineering. However, deep learning models typically require large amounts of labeled data and significant computational resources for training.

2.2.3 Convolutional neural networks (CNNS)

CNNs are a class of deep neural networks specifically designed for processing grid-like data, such as images. They consist of multiple layers, including convolutional, pooling, and fully connected layers, which extract progressively abstract features from input images. CNNs have shown remarkable performance in various computer vision tasks, including crack detection, by automatically learning and capturing complex visual patterns within crack images. The hierarchical structure of CNNs allows them to identify local features as well as global patterns, making them a suitable choice for crack identification and classification.

Advantages of CNNs in Crack Detection Hierarchical Feature Learning: CNNs can automatically learn and represent hierarchical features within images, capturing intricate patterns and textures characteristic of cracks. Local and Global Context: CNNs can simultaneously capture local and global contextual information, aiding in accurate crack identification. Robustness: CNNs are robust to variations in lighting, orientation, and noise, making them suitable for real-world crack detection scenarios. Disadvantages of CNNs in Crack Detection:

Data Requirements: CNNs require a large amount of labeled training data to generalize well to diverse crack patterns and

TABLE 4 Papers with their own datasets.

Paper	Their own dataset
Li et al. (2022b)	This study employed image data obtained from CiCS, a rapid road condition detection system in China. The dataset consisted of 1,923 pavement cracking images in "jpg" format, each with a resolution of 96 dpi (3,024 × 1887 pixels). Out of these, 1,538 images were allocated for training, while the remaining 385 were designated for testing
Yang et al. (2022)	The dataset comprised 10,400 images acquired by a vehicle equipped with a highway condition monitor, encompassing a total of 202,840 labeled instances of pavement distress
Lee et al. (2022)	The dataset consisted of 4,000 images, equally divided into 2,000 negative and 2,000 positive samples
Mo et al. (2022)	The hazardous clearance set comprised 1,200 samples, while the non-hazardous clear-ance set contained 800 samples
Li et al. (2022c)	The dataset consisted of 600 images captured using the Teledyne Dalsa S3-24-02k40, a high-speed linear array industrial digital camera
Zhang et al. (2023b)	The dataset comprises approximately 800 photos of bridge cracks, collected manually using a digital camera with a resolution of 1,664 × 1,664 pixels in JPG format. In total, 4,033 images were labeled, capturing various weather conditions
Ha et al. (2022)	Total 6,650 images
Maslan and Cicmanec (2023)	Total 3,279 images
Loverdos and Sarhosis (2022)	Total 107 fully annotated images of masonry structures
Wibowo et al. (2022)	Total 2,516 images obtained through web scraping
Wang (2023)	The dataset initially consisted of 284 images, each with a resolution of 1706 × 1,280 pixels and 8 bits per channel in RGB format, captured on the campus of Hefei University using a smartphone. After applying augmentation techniques, the dataset expanded to encompass 6,000 images
Xu et al. (2022)	The dataset comprised 148 images of pavement cracks, which were captured using smartphones and covered various types of cracks. Additionally, data augmentation techniques were applied to expand the dataset by 50%
Kapadia et al. (2023)	The dataset was meticulously curated, containing over 80,000 images of cracks and an equal number of non-crack images sourced from concrete cube surfaces
Ma et al. (2022b)	The MaDataSet, a collection of 474 images depicting cracks in ancient timber structures, was compiled from the Bawang Academy at Shen-yang Jianzhu University
Wan et al. (2022)	The dataset comprised more than 300 crack images from the Jing-Hang Grand Canal Extra Large Bridge site and over 1,500 crack images from the laboratory. Following data augmentation, the total image count expanded to 7,000
Ren et al. (2022)	A high-resolution camera, affixed to a vehicle, gathered 9,650 images portraying diverse pavement cracks for the dataset.
Elghaish et al. (2022)	The dataset comprised 4,663 images of highway cracks, categorized into three groups "vertical cracks," "horizontal and vertical cracks," and "diagonal cracks."
Chu et al. (2023)	Total 6,000 images from Pakistani roads
Wu et al. (2022)	125 images of cracked concrete with dimensions measuring 3120 × 4160 pixels
Nomura et al. (2022)	Totally around 100,000 images
Yu et al. (2022)	437 RGB images with their segmentation labels
Yu et al. (2022)	The dataset comprises 1,000 digital photos captured using a Canon EOS 5DS R camera during regular bridge inspections. Among these photos, 487 contained cracks, and through augmentation techniques, this number was expanded to 3,365 images
Munawar et al. (2022a)	The benchmarking dataset included 600 images depicting damage to buildings located in Sydney
Kun et al. (2022)	The dataset consisted of 385 images of bridges taken at Zhongshan Bridge in Gansu China, using an I-800 Airborne Plane Array Camera
Kolappan Geetha et al. (2023)	The dataset used to train the 1D DFT-CNN model consisted of 1,492 images containing cracks and 1,321 images without cracks on concrete surfaces, encompassing diverse optical conditions
Hammouch et al. (2022)	The dataset was generated from video frames supplied by the Moroccan National Center for Road Studies and Research (CNER) through the SMAC system, leading to a collection of 3,287 images that were manually annotated
Deng et al. (2023)	Total 6,840 images
Tan and Dong (2023)	The dataset was thoughtfully curated to include 2000 images, primarily due to hardware (GPU) constraints. Segmentation masks were generated using LabelImgPlus
Paramanandham et al. (2023)	The authors utilized their own dataset for their research

(Continued on following page)

TABLE 4 (*Continued*) Papers with their own datasets.

Paper	Their own dataset
Lee and Huh (2022)	Total 800 multisensory images
Kim et al. (2022)	The wall quality dataset consisted of 5,000 images showcasing various defects such as cracks, holes, efflorescence, damp patches, and spall issues, with 1,000 images dedicated to each category
Kou et al. (2022)	The dataset encompassed 380 images of rail cracks, gathered over a year in diverse weather, traffic, and wear conditions
Yuan et al. (2022)	300 self-captured pavement crack images
Liu et al. (2023)	Total 21,547 images
Guo et al. (2023)	Total 88 images
Li et al. (2023b)	The dataset was generated from drone-captured rock mass crack images, resulting in a VOC dataset that employed advanced data augmentation techniques
Kim et al. (2023a)	The image dataset was sourced from the Korea Expressway Corporation's bridge monitoring system, consisting of a total of 192 images
Tse et al. (2023)	The dataset included 4,000 crack images, covering nine distinct crack types with different orientations
Lee and Yoo (2023)	The crack dataset consisted of 11,226 image sets along with corresponding masks, facilitating precise crack detection and non-semantic object removal
Quqa et al. (2022)	A dataset comprising images of welding joints from a long-span steel bridge, captured using high-resolution consumer-grade digital cameras
Ji et al. (2022)	Total 11,449 images, including 4,650 pavement and 6,799 concrete images
Chen et al. (2022)	Total 1,452 images

TABLE 5 Papers with Public dataset and their own dataset.

Paper	Public dataset + their own dataset
Ashraf et al. (2023)	This study utilized data from two sources: RDD2022, a publicly available online dataset, and a second set of data collected from the roads of Malaysia
Yong and Wang (2022)	727 images by 4-megapixel Hikvision industrial camera MV-CA060–10 GC and Crack500 dataset
Lee et al. (2023)	A total of 1,235 test images were obtained from drone footage on the sea bridge, in addition to 10,789 open images sourced from datasets including Crack Forest Dataset, Crack500, CrackTree200, Deep Crack, and GAPs384
Zhao et al. (2023b)	UAVRoadCrack, and public datasets: CRKWH100, and CrackLS315 Total 7,403 images
Munawar et al. (2022b)	All the self-built dataset images were acquired using the DJI-M200 UAV, which utilizes vertical take-off and landing (VTOL) technology. This dataset amalgamated data from Crack Forest Dataset (CFD), Crack500, and GAPs, resulting in a comprehensive collection of 1,300 images
Kang and Cha (2022)	A total of 1,203 images were included in the dataset, and they underwent extensive augmentation using synthesis techniques. Additionally, the dataset was evaluated using 545 testing images sourced from both existing datasets and their proprietary data
Panta et al. (2023)	The dataset of levee crack images was gathered over several years by field inspectors from the New Orleans district of the U.S. Army Corps of Engineers (USACE). Initially, it contained 1,650 images. To enhance the dataset, 101 additional levee crack images were annotated using the VGG Image Annotator tool. For comprehensive model analysis, the DeepCrack road crack dataset was also utilized. This dataset's test subset comprised 237 images, each accompanied by its respective masks
Inam et al. (2023)	Self-built dataset from Pakistan and the SDNET2018 dataset
Kao et al. (2023)	Crack images were sourced from various devices, including smartphones, a camera mounted on a UAV (Unmanned Aerial Vehicle), and the open-source deep-learning dataset SDNET2018
Li et al. (2022a)	A total of 48,000 images were used, including those generated through augmentation (from a self-created dataset), and the Crack Forest Dataset (CFD) and GAPs images were included for evaluation

variations. Computational Intensity: Training and fine-tuning CNNs can be computationally intensive, necessitating powerful hardware and resources. Interpretability: The inner workings of CNNs can be challenging to interpret, limiting their explainability in critical applications. In the 85 articles surveyed, a diverse array of algorithms and methods were employed for crack detection. Notably, the following algorithms emerged as prominent choices, showcasing their effectiveness in addressing the intricacies of crack identification and classification:

CNN (Convolutional Neural Networks): CNNs, as previously discussed, are widely used for crack detection due to their ability to capture complex patterns and textures in images. They have been

TABLE 6 Main Algorithms with their related papers.

Algorithm	Paper
CNN	Yang et al. (2023), Chu et al. (2023), Inácio et al. (2023), de León et al. (2023), Guo et al. (2023), Kim et al. (2023a), Lee and Yoo (2023), Jayaraju et al. (2022), Lee et al. (2022), Yong and Wang (2022), Wu et al. (2022), Lu et al. (2022), Quqa et al. (2022), Chen et al. (2022), Yadav et al. (2022), Golding et al. (2022), Wang et al. (2022a)
Deep Learning	Golding et al. (2022), Kou et al. (2022), Kun et al. (2022), Liu et al. (2022), Pu et al. (2022), Wan et al. (2022), Kolappan Geetha et al. (2023), Liu et al. (2023), Ngo et al. (2023), Popli et al. (2023), Shim et al. (2023), Yang et al. (2023)
YOLO	Ashraf et al. (2023), Zhang et al. (2023b), Maslan and Cicmanec (2023), Gooda et al. (2023), Deng et al. (2023), Yu and Zhou (2023), Inam et al. (2023), Li et al. (2023b), Tse et al. (2023), Kao et al. (2023), Xu et al. (2022), Li et al. (2022b), Yang et al. (2022), Ma et al. (2022b), Ren et al. (2022), Nomura et al. (2022), Yu et al. (2022)
UNet	Kim et al. (2023b), Gooda et al. (2023), Bai et al. (2023), Deng et al. (2023), Inam et al. (2023), Zhao et al. (2023a), Li et al. (2023a), Wang et al. (2022b), Li et al. (2022c), Ha et al. (2022), Loverdos and Sarhosis (2022), Mohammed et al. (2022), Jing et al. (2022), Ji et al. (2022), Wang et al. (2022a)
ResNet	Wang et al. (2022a), Islam et al. (2022), Ji et al. (2022), Kim et al. (2022), Paramanandham et al. (2022), Siriborvornratanakul (2022), Wibowo et al. (2022), Li et al. (2023a), Bai et al. (2023), Deng et al. (2023), Paramanandham et al. (2023), Qayyum et al. (2023)
Rsef	Kim et al. (2023b)
Ensemble Learning	Lee et al. (2023)
CrackNet	Gharehbaghi et al. (2022), Zhao et al. (2023b)
Mask RCNN	Lv et al. (2023), Xu et al. (2022)
Fast RCNN CrackSN	Mo et al. (2022), Xu et al. (2022), Zhao et al. (2022) Wang (2023)
Inceptionv3	Kapadia et al. (2023), Paramanandham et al. (2023), Qayyum et al. (2023)
IterLUNet	Panta et al. (2023)
VGG	Wang et al. (2022a), Elghaish et al. (2022), Hammouch et al. (2022), Islam et al. (2022), Ji et al. (2022), Nomura et al. (2022), Paramanandham et al. (2022), Wibowo et al. (2022), Zhang et al. (2022), Guo et al. (2023), Paramanandham et al. (2023), Philip et al. (2023)
MobileNet	Ha et al. (2022), Philip et al. (2023), Qayyum et al. (2023)
Xception	Philip et al. (2023)
GoogleNet	Qayyum et al. (2023), Elghaish et al. (2022)
ShuffleNet	Qayyum et al. (2023)
Omni-Dimensional Dynamic Convolution	Tan and Dong (2023)
PIRM	Paramanandham et al. (2023)
CTCD-Net	Zhang et al. (2023a)
DenseNet	Li et al. (2022a), Wang et al. (2022a), Islam et al. (2022)
AlexNet	Elghaish et al. (2022), Islam et al. (2022), Paramanandham et al. (2022)
SqueezeNet	Ha et al. (2022), Wang (2023)
DeepLab	Loverdos and Sarhosis (2022), Wang et al. (2022a), Siriborvornratanakul (2022), Yu et al. (2022)
LinkNet	Loverdos and Sarhosis (2022)
FPN	Loverdos and Sarhosis (2022)
ViT	Ali et al. (2022)
Adversarial Network	Munawar et al. (2022a), Yuan et al. (2022)
STRNet	Kang and Cha (2022)

applied in various architectures and configurations to achieve high accuracy in crack identification.

Deep Learning: Deep learning approaches beyond CNNs have been leveraged to enhance crack detection:

YOLO (You Only Look Once): YOLO is a real-time object detection algorithm that can identify and locate multiple objects in an image simultaneously. It has been adapted for crack detection to provide efficient and accurate localization of cracks within images.

TABLE 7 Algorithms used in 85 articles.

#	Paper	Algorithm
1	Ashraf et al. (2023)	Custom YOLOv7
2	Yang et al. (2023)	Deep CNN
3	Kim et al. (2023b)	Rsef based on U-net namely, Rsef-Edge
4	Lee et al. (2023)	Ensemble learning
5	Zhang et al. (2023b)	Yolo v4
6	Zhao et al. (2023b)	CrackNet
7	Maslan and Cicmanec (2023)	YOLO v2
8	Lv et al. (2023)	Mask R-CNN
9	Wang (2023)	CrackSN built on the Adam-SqueezeNet architecture
10	Gooda et al. (2023)	EfficientNet with residual U-Net for segmentation, YOLO v5 for crack detection
11	Kapadia et al. (2023)	The Inceptionv3 model
12	Ngo et al. (2023)	Deep learning
13	Chu et al. (2023)	CNN
14	Bai et al. (2023)	ResNet and ResNet + UNet
15	Kolappan Geetha et al. (2023)	Deep Learning
16	Panta et al. (2023)	Iterative Loop UNet (IterLUNet)
17	Philip et al. (2023)	VGG16, VGG19, ResNet 50, MobileNet, and Xception
18	Qayyum et al. (2023)	GoogLeNet, MobileNet-V2, Inception-V3, ResNet18, ResNet50, ResNet101, and ShuffleNet
19	Inácio et al. (2023)	Multi-class CNN
20	Deng et al. (2023)	YOLOv5 crack detection and Res-UNet segmentation
21	Tan and Dong (2023)	A pyramidal residual network, employing an encoder-decoder architecture, incorporates Omni Dimensional Dynamic Convolution for its operations
22	Paramanandham et al. (2023)	The Pixel-Intensity Resemblance Measurement (PIRM) rule was applied in conjunction with VGG-16, ResNet-50, and InceptionResNet-V2 models for the purpose of crack detection
23	Yu and Zhou (2023)	A novel approach for crack detection, named YOLOv5-CBoT, is introduced by enhanc-ing the YOLOv5 network with a Bottleneck Transformer
24	Zhang et al. (2023a)	CTCD-Net: A Cross-layer Transmission network for tiny road Crack Detection
25	de León et al. (2023)	A novel crack segmentation algorithm has been developed, which combines the theory of minimal path selection with a region-based approach. This method involves the segmentation of texture features extracted using Gabor filters
26	Inam et al. (2023)	YOLOv5 for crack detection and U-Net for segmentation
27	Liu et al. (2023)	Deep convolutional network (Single Shot MultiBox Detector (SSD))
28	Guo et al. (2023)	Adopted convolutional neural network (CNN) (VGG16 + Focal Loss)
29	Li et al. (2023b)	The YOLOv7 with attention mechanism
30	Kim et al. (2023a)	CNN
31	Tse et al. (2023)	Improved YOLOv4 with an attention module
32	Kao et al. (2023)	YOLOv4
33	Lee and Yoo (2023)	Fast encoder-decoder network with scaling attention
34	Zhao et al. (2023a)	U-Net
35	Shim et al. (2023)	A novel deep neural network has been introduced, accompanied by an adversarial learning-based balanced ensemble discriminator network
36	Li et al. (2023a)	Segmentation by ResNet50 as a UNet model

(Continued on following page)

TABLE 7 (*Continued*) Algorithms used in 85 articles.

#	Paper	Algorithm
37	Popli et al. (2023)	Deep learning
38	Xu et al. (2022)	Fast RCNN, Mask RCNN and compare with YOLO.
39	Jayaraju et al. (2022)	CNN
40	Zhang et al. (2022)	FPN-vgg16
41	Wang et al. (2022b)	U-Net and the dual-attention network (DANet), and efficient mobile-attention X-network (MA-Xnet)
42	Li et al. (2022b)	YOLOv4-3
43	Yang et al. (2022)	YOLOv5s
44	Lee et al. (2022)	CNN
45	Mo et al. (2022)	Fast R-CNN
46	Yong and Wang (2022)	An end-to-end real-time pavement crack segmentation network, denoted as RIIAnet, has been developed
47	Li et al. (2022c)	U-Net and a side-output part: SoUNet
48	Islam et al. (2022)	VGG16, ResNet18, DenseNet161, and AlexNet
49	Ha et al. (2022)	SqueezeNet, U-Net, and Mobilenet-SSD
50	Loverdos and Sarhosis (2022)	U-Net, DeepLabV3+, U-Net (SM), LinkNet (SM), and FPN (SM)
51	Ali et al. (2022)	The vision-transformer (ViT)
52	Wibowo et al. (2022)	VGG16 and RestNET50
53	Pu et al. (2022)	Deep convolutional neural network (DCNN)
54	Munawar et al. (2022b)	Cycle generative adversarial network (CycleGAN)
55	Ma et al. (2022b)	YOLO v3, YOLO v4s-mish, and YOLO v5s
56	Wan et al. (2022)	Deep Learning
57	Ren et al. (2022)	YOLOV5
58	Kang and Cha (2022)	Semantic transformer representation network (STRNet)
59	Siriborvornratanakul (2022)	DeepLabV3-ResNet101
60	Elghaish et al. (2022)	AlexNet, VGG16, VGG19, GoogleNet.
61	Wu et al. (2022)	Full convolutional neural networks FCN-8s, FCN-16s, and FCN-32s
62	Liu et al. (2022)	Deep Domain Adaptation-based Crack Detection Network (DDACDN)
63	Nomura et al. (2022)	YOLOv2 + VGG16
64	Yu et al. (2022)	DeepLabV3+
65	Yu et al. (2022)	YOLOv5
66	Munawar et al. (2022a)	CNN and a cycle generative adversarial network (CycleGAN)
67	Kun et al. (2022)	Deep bridge crack classification (DBCC)-Net
68	Mohammed et al. (2022)	U-Net
69	Hammouch et al. (2022)	VGG-19
70	Lee and Huh (2022)	This study introduces a dataset.
71	Lu et al. (2022)	Multi-scale crack detection network: MSCNet
72	Kim et al. (2022)	Conv2D ResNet.
73	Kou et al. (2022)	Deep Learning
74	Zhao et al. (2022)	Fast R-CNN

(Continued on following page)

TABLE 7 (*Continued*) Algorithms used in 85 articles.

#	Paper	Algorithm
75	Jing et al. (2022)	AR-UNet based on UNet
76	Gharehbaghi et al. (2022)	FastCrackNet
77	Yuan et al. (2022)	Generative adversarial network
78	Paramanandham et al. (2022)	Alexnet, VGG16, VGG19 and ResNet-50
79	Quqa et al. (2022)	CNN
80	Ji et al. (2022)	U-Net. ResUNet, VGGU-Net, and EfficientU-Net.
81	Chen et al. (2022)	Enhanced graph network branch
82	Yadav et al. (2022)	Multi-scale feature fusion (3SCNet + LBP + SLIC)
83	Golding et al. (2022)	Deep learning-based autonomous crack detection method using CNN
84	Li et al. (2022a)	Dense boundary refinement network (DBR-Net)
85	Wang et al. (2022a)	Several semantic segmentation models were explored, including Fully Convolutional Network (FCN), Global Convolutional Network (GCN), Pyramid Scene Parsing Net- work (PSPNet), UPerNet, and DeepLabv3+. These models were coupled with various backbone architectures, including VGG, ResNet, and DenseNet, to investigate their performance

UNet: UNet is a convolutional neural network architecture designed for biomedical image segmentation. Its U-shaped architecture enables precise segmentation of crack regions, making it well-suited for crack detection and localization.

ResNet (Residual Network): ResNet is a deep convolutional neural network architecture known for its ability to mitigate the vanishing gradient problem in deep networks. ResNet-based models have been effective in capturing intricate features within crack images.

Rsef: Residual Network with Feature Shrinking (Rsef) is a variant of ResNet that incorporates feature shrinking to reduce the computational complexity of the network. It has been used for efficient and accurate crack detection.

Ensemble Learning: Ensemble learning techniques, such as combining predictions from multiple models, have been employed to enhance crack detection accuracy and robustness, demonstrating improved performance over individual models.

CrackNet: CrackNet is a specialized architecture designed explicitly for crack detection. It employs convolutional and pooling layers to capture crack patterns and structural features, resulting in high accuracy.

Mask RCNN and Fast RCNN: These architectures extend CNNs to perform instance segmentation, enabling accurate identification and localization of individual cracks within images.

Inceptionv3, IterLUNet, VGG, MobileNet, Xception, GoogleNet, ShuffleNet, and Omni-Dimensional Dynamic Convolution: These deep learning architectures have been explored to optimize feature extraction and crack detection performance, leveraging their unique design principles.

Pixel-intensity resemblance measurement (PIRM), CTCD-Net, and DeepLab: Specialized techniques have been developed to assess pixel-level resemblance and semantic segmentation, allowing for detailed and fine-grained crack detection.

Adversarial Network and STRNet: Adversarial networks and architecture variants like STRNet have been used to enhance model robustness and generalization, contributing to more reliable crack detection. In Table 6 we can see the main methods which mentioned above, with the papers that they use them. The comprehensive details and utilization of these algorithms in the surveyed papers can be found in Table 7 providing valuable insights into their specific applications and performance in crack detection tasks.

3 Results and discussion

3.1 Results

The "Results" section of a crack detection study is a critical component that showcases the performance and effectiveness of the proposed methodologies. It provides a quantitative assessment of how well the developed algorithms and models perform in detecting and classifying cracks in various structures. This section serves as a validation of the proposed solutions, allowing researchers to evaluate their contributions and compare them to existing methods.

3.2 Metrics for evaluating results

To objectively evaluate the performance of crack detection algorithms, researchers employ a variety of metrics that assess different aspects of model performance. These metrics provide insights into the accuracy, precision, recall, and overall effectiveness of the methods. Let's explore some of the commonly used metrics in crack detection research:

Accuracy: Accuracy measures the proportion of correctly predicted crack and non-crack instances among all predictions. It provides an overall assessment of the model's correctness but might be skewed in imbalanced datasets.

Precision: Precision determines the proportion of accurate positive forecasts to all instances of positive predictions. It indicates the model's ability to correctly identify positive cases, minimizing false positives.

Recall (Sensitivity): The ratio of accurate positive predictions to all actual positive cases is calculated using recall. It highlights the

model's capacity to identify all positive cases, minimizing false negatives.

F1 Score: The harmonic mean of recall and precision is the F1 score. It balances the trade-off between precision and recall, providing a single metric to assess the model's performance.

Intersection over Union (IoU): IoU calculates how much of the expected and actual bounding boxes or masks overlap. It is commonly used in object detection and segmentation tasks.

Mean Average Precision (mAP): The average precision across various confidence threshold levels is determined by mAP. It is often used in object detection tasks to evaluate the precision-recall curve.

Receiver Operating Characteristic (ROC) Curve: The true positive rate *versus* the false positive rate at different categorization criteria are plotted on the ROC curve. A popular statistic used to evaluate the effectiveness of models is the area under the ROC curve (AUC-ROC).

Precision-Recall (PR) Curve: The PR curve illustrates the trade-off between the two metrics by plotting recall against precision at various categorization levels.

Dice Coefficient: The Dice coefficient measures the similarity between the predicted and ground truth segmentation masks.

Matthews Correlation Coefficient (MCC): MCC offers a balanced statistic for binary classification tasks by accounting for true positive, true negative, false positive, and false negative predictions.

These metrics collectively offer a comprehensive view of a crack detection model's performance. Researchers select the appropriate metrics based on the specific objectives of their study and the nature of the crack detection problem.

Supplementary Table S1 presents a detailed overview of the results obtained from the 85 reviewed papers. The table offers insights into the performance of various crack detection algorithms across different metrics, providing a comprehensive analysis of their effectiveness in real-world scenarios. The subsequent sections delve into the specific findings and trends observed in the evaluated papers, shedding light on the advancements and challenges in crack detection technologies.

3.3 Discussion

The field of crack detection has witnessed significant advancements owing to the integration of deep learning techniques. In this discussion, we delve into the collective insights, contributions, and limitations presented across a diverse range of recent research papers. As we navigate through the reviewed studies, we categorize them based on their innovative approaches and the identified gaps or limitations.

The introduction of the "Custom YOLOv7" model by Ashraf et al. (2023) marks a substantial stride in crack detection. This model achieves exceptional accuracy on both the RDD2022 dataset and a custom dataset. While the model's performance is remarkable, opportunities lie in refining its efficiency and exploring pixel-level segmentation strategies. Yang et al. (2023)'s "AttentionCrack" network presents a promising solution to enhance crack detection accuracy by addressing inaccuracies in boundary localization. The model demonstrates impressive F1 scores on benchmark datasets. However, the authors highlight potential areas of exploration, such as attention mechanisms and dilated convolution modules, to further enhance performance. Kim Y. et al. (2023)'s "Rsef-Edge," built upon the U-net architecture, stands out for achieving an accuracy rate of 97.36%. The paper suggests the implementation of an edge computing-based crack detection system. Nevertheless, challenges and potential advantages related to distributed deep learning form an essential part of the ongoing discussion. The "Stacking Ensemble Model" proposed by Lee et al. (2023) offers a novel approach to crack segmentation by leveraging ensemble learning. This model achieves an Intersection over Union (IoU) of 0.74, significantly outperforming FCN-8s. The focus on stacking ensemble learning and its impact on performance opens avenues for further investigation. Zhang J. et al. (2023)'s "Automated Yolo v4" introduces a method that emphasizes precision, recall, and F1 scores, showcasing a compelling alternative to existing approaches. The paper highlights the model's efficiency and compactness, making it a viable solution. However, addressing the challenges posed by imbalanced data remains a crucial direction for future research.

In "CrackNet" and "CrackClassification," Zhao Y. et al. (2023) contribute with their novel CrackNet model and CrackClassification algorithm. The study reports average precision (AP) scores for the CrackNet network across various datasets. The insights from this work shed light on the potential of the proposed methods in the context of crack detection. Maslan and Cicmanec (2023) propose the utilization of Yolo v2 for crack detection, resulting in an average precision (AP) of 0.89. This work showcases the model's competency in crack detection and sets the stage for discussions on the selection of YOLO versions for optimal results. Lv et al. (2023)'s "Mask R-CNN" presents a robust solution based on the mask region-based Convolutional Neural Network. The model achieves accuracy rates ranging from 95% to 99% on diverse datasets. While the model's performance is commendable, the paper also acknowledges the need for thorough comparative analysis and the selection of pooling layers. Wang (2023)'s "CrackSN" system, built on the Adam-SqueezeNet architecture, achieves an accuracy of 97.3% in classifying cracked patches. The authors discuss the positive aspects and limitations of their system, including its reliance on specific datasets and the potential for improving pixel-level accuracy. The novel proposal of "EfficientNet with Residual U-Net" by Gooda et al. (2023) combines segmentation and detection techniques to achieve an impressive accuracy of 99.35%. The paper's methodology and results provide a strong foundation for further exploration, while the discussion raises questions about computational requirements and improvements in the proposed methods. Kapadia et al. (2023)'s work on the "Inceptionv3" model adds valuable insights into accuracy, cross-entropy, precision, recall, and F-score values. The study acknowledges the challenges posed by acquired images and underlines the limitations of conventional algorithms in the domain of crack detection. Ngo et al. (2023)'s deep learning approach showcases an accuracy of 95.19% for crack detection. This work accentuates the importance of reliable datasets and addresses limitations in previous crack detection methods. The study's emphasis on dataset quality sets the stage for further investigation.

Chu et al. (2023)'s "Pothole Crack Detection (PCD)" model leverages a CNN-based approach to achieve remarkable precision and recall rates. The paper introduces a novel deep learning method that extends beyond crack detection to address road damage and pothole identification. The emphasis on decision support systems and a self-collected dataset enhances the practical relevance of the work. Bai et al. (2023)'s proposal to employ ResNet and ResNet + UNet for crack detection results in an accuracy of 67.6%. While the paper highlights the potential of these architectures, it also acknowledges the need for more labeled images and explores the utilization of benchmark datasets. The discussion reflects the ongoing pursuit of accurate and efficient crack detection solutions. Kolappan Geetha et al. (2023) take an innovative approach by employing an iterative differential sliding-window-based local image processing technique for missing crack detection. The study's focus on enhancing efficiency and introducing a novel scheme for eliminating missing shallow propagating crack segments offers new avenues for further research.

Inam et al. (2023)'s integration of YOLOv5 and U-Net for bridge crack detection demonstrates the potential of combining detection and segmentation approaches. This novel combination contributes to the field by showcasing the advantages of leveraging both models in tandem. The study also raises considerations for applying this approach to bridge crack detection in developing countries. Lee et al. (2023)'s "Image Processing and Deep Learning" method introduces a deep convolutional network (SSD) for object detection in tunnel images. The study compares various CNN models based on accuracy and discusses challenges in implementation and real-world feasibility. While the method holds promise, the paper acknowledges the need for more in-depth discussion on implementation challenges. Guo et al. (2023)'s adoption of a CNN (VGG16 + Focal Loss) for crack detection and quantification presents a promising way to estimate defect dimensions on complex structures. The paper's validation through gauge measurements and point cloud data opens avenues for applying the proposed approach to diverse scenarios. Li et al. (2023b)'s proposal of YOLOv7 with an attention mechanism for crack detection showcases improvements in precision and recall rates. The model's superior performance adds to the ongoing discourse on achieving a balance between accuracy and inference speed. Kim J.-Y. et al. (2023)'s "Blurred and Indistinct Concrete Crack Detection Framework" introduces a framework for detecting challenging blurred and indistinct concrete cracks. The paper explores the effectiveness of CNN models like AlexNet, VGG-16, and ResNet152 in classification and highlights the limitations of image filtering and thresholding methods. This work emphasizes the importance of tackling complex scenarios in crack detection. Tse et al. (2023)'s "Improved YOLOv4 with Attention Module" showcases an enhanced YOLOv4 model with an attention module that achieves high mean average precision (mAP). The study's focus on improving model efficiency and performance underscores the dynamic nature of crack detection research. Kao et al. (2023)'s "Combining YOLOv4 for Crack Detection" presents an approach utilizing YOLOv4 for accurate crack detection, validated through quantitative crack test methodologies. This work emphasizes the significance of image processing and edge detection techniques in achieving reliable results.

Lee and Yoo (2023)'s "Fast Encoder-Decoder Network with Scaling Attention" contributes a fast encoder-decoder network with scaling attention to the field. The model's competitive results and focus on detecting fine-grained cracks point towards the ongoing efforts to balance computational efficiency and precision. Zhao F. et al. (2023)'s "U-Net-Based Crack Segmentation with Morphological Network" introduces a novel crack segmentation method employing a U-Net-based architecture with a morphological network and multi-loss function. The proposed method's capability to improve crack segmentation performance under polarized light conditions adds a nuanced perspective to the field. Shim et al. (2023)'s "Stereo Adversarial Learning-Based Balanced Ensemble Discriminator Network" unveils a novel deep neural network with an adversarial learning-based balanced ensemble discriminator network. The model's performance in terms of intersection-over-union and F1 scores presents an intriguing avenue for addressing challenges posed by varying environmental conditions. Li et al. (2023a)'s "Intelligent Deep Learning for Crack Feature Extraction and Segmentation" introduces a two-stage transfer learning approach using ResNet50 and multilayer parallel residual attention (MPR) for crack feature extraction and segmentation. The study's emphasis on improvements over the benchmark UNet model underscores the potential of incorporating advanced neural network architectures. Popli et al. (2023)'s integration of a robot vision system with deep learning for road crack identification culminates in the identification of Xception as the most accurate and predictive model among the tested algorithms. The study's call for comprehensive investigations into crack detection complexities highlights the multifaceted nature of real-world applications. Xu et al. (2022)'s comparison of Fast RCNN, Mask RCNN, and YOLO for crack detection brings forth insights into the performance of these models. While Fast RCNN emerges with better results, this paper illustrates the importance of understanding the trade-offs between different detection architectures. Jayaraju et al. (2022)'s CNN-based approach for high-accuracy crack detection in building structures offers an efficient and objective solution. The paper's focus on utilizing a large dataset and CNN for precise detection draws attention to the potential of data-driven approaches in enhancing accuracy. Zhang et al. (2022)'s proposal for crack detection in earthen heritage sites using FPN-vgg16 combines effective crack extraction and transfer learning. The study's engagement with challenges related to deployment and uncertainty in crack attributes underscores the nuanced considerations in heritage preservation.

Wang et al. (2022b)'s "MA-Xnet" introduces an efficient mobile-attention X-network for crack detection. While the paper celebrates state-of-the-art performance and attention mechanisms, it acknowledges the need for further exploration in dataset generalization and computational complexity analysis. Li L. et al. (2022)'s utilization of Conv2D ResNet with an exponential activation layer yields superior results in wall defect classification. The study's call for further validation and assessment across different convolutional layers and loss functions underscores the iterative nature of deep learning research. Islam et al. (2022)'s "CNN-Based Transfer Learning for Crack Detection" introduces a transfer learning approach based on CNN for robust crack detection. The paper's demonstration of high accuracy across various deep learning models accentuates the importance of

model selection in achieving reliable results. The need for diverse datasets and exploration of alternative neural network architectures remains open for further investigation. Ha et al. (2022)'s assessment of SqueezeNet, U-Net, and Mobilenet-SSD models for crack assessment highlights their high accuracy in defect classification. The paper's emphasis on accurate severity assessment and limitations involving depth information and system size draw attention to the complexity of evaluation metrics in real-world applications.

Loverdos and Sarhosis (2022)'s comparison of U-Net, DeepLabV3+, LinkNet (SM), and FPN (SM) models underscores their high accuracy in crack detection. The positive outcomes achieved by the block-detection model and crack detection model bring to light the significance of model selection and its impact on accuracy. Ali et al. (2022)'s vision-transformer (ViT) classifier for crack classification, localization, and segmentation reflects a promising integration of advanced algorithms. The high accuracy, precision, recall, and F1 scores achieved through this integration affirm the potential of combining state-of-the-art techniques. Wibowo et al. (2022)'s utilization of transfer learning with VGG16 and ResNet50, combined with ANN and kNN, in wall crack classification showcases a fusion of methodologies for enhanced accuracy. The paper's recognition of dataset quality and variety serves as a reminder of the fundamental role data plays in the efficacy of deep learning models. Pu et al. (2022)'s employment of a deep convolutional neural network (DCNN) with an encoder-decoder module for semantic segmentation and classification accentuates the significance of accuracy improvement. The promising outcomes demonstrated underscore the iterative nature of model enhancement and the potential of deep learning techniques. Munawar et al. (2022a)'s investigation into crack detection using a modified deep hierarchical CNN architecture and CycleGAN underscores the utility of guided filtering and CRFs for pixel-wise segmentation. The exploration of various accuracy metrics and techniques emphasizes the multifaceted nature of crack detection research. Ma J. et al. (2022)'s comparative evaluation of YOLO v3, YOLO v4s-mish, and YOLO v5s for crack detection in ancient timber structures provides insights into the strengths of different architectures. While YOLO v3 emerges as a strong performer, the study's focus on training speed speaks to the ongoing pursuit of efficient and accurate detection methods.

Wan et al. (2022)'s combination of SSD and an eight-neighborhood algorithm demonstrates high precision and recall in crack detection. The paper's recognition of challenges in length and width identification and its reference to specific scenarios highlight the diverse environments in which crack detection operates. Ren et al. (2022)'s utilization of YOLOv5 for precise pavement crack detection showcases advancements in model accuracy. The proposed method's ability to improve detection performance over existing methods reiterates the iterative nature of model development. Kang and Cha (2022)'s introduction of STRNet, a semantic transformer representation network, achieves high precision, recall, F1 score, and mIoU in crack segmentation. The paper's exploration of false positives and negatives underscores the complexities of segmenting intricate crack patterns. Siriborvornratanakul (2022)'s adoption of DeepLabV3-ResNet101 for damage detection addresses complex scene detection using deep learning solutions. The paper's identification of gaps in pixel-level localization highlights the need for holistic crack detection methodologies. Elghaish et al. (2022)'s development of a new CNN model that outperforms pre-trained models presents an exciting avenue for infrastructure maintenance. The call for ongoing investigations serves as a reminder of the evolving nature of crack detection research.

Wu et al. (2022)'s exploration of FCN architectures for crack detection showcases the ongoing pursuit of improving accuracy and handling complex crack patterns. The challenges posed by factors like illumination and the desire for precise part detection underscore the dynamic nature of detection techniques. Liu et al. (2022)'s incorporation of domain adaptation into DDACDN for crack detection highlights the model's high accuracy. The call for quantitative evaluation, active learning, and consideration of multi-scale objects acknowledges the intricacies of real-world implementation. Nomura et al. (2022)'s evaluation of YOLOv2 + VGG16 for damage detection emphasizes the importance of improving recall and addressing challenges related to over-detection. The paper's recognition of the need for automating detection processes aligns with the drive for efficiency in detection methodologies. Yu et al. (2022)'s contributions to intelligent performance improvements in DeepLabV3+ and YOLOv5 demonstrate the potential of these techniques in various contexts. The discussion of dataset scaling, new loss functions, and filtering methods invites further exploration and refinement. Munawar et al. (2022b)'s exploration of a CNN architecture coupled with CycleGAN for crack detection showcases the potential of guided filtering and offers insights into global accuracy, class average accuracy, intersection of union, precision, recall, and F-score metrics. The focus on CNN architecture and CycleGAN exemplifies the synergy between different techniques in enhancing crack detection. Kun et al. (2022)'s Deep Bridge Crack Classification (DBCC)-Net presents a unique approach by converting target detection from regression to binary classification. The paper's emphasis on achieving higher Miou while acknowledging limitations in reasoning time and available research data underscores the importance of innovative strategies. Mohammed et al. (2022)'s semi-supervised learning model for crack detection provides an avenue for reducing the need for labeled data while maintaining accuracy. The paper's alignment with efficient data utilization and training time optimization contributes to the ongoing exploration of deep learning techniques. Hammouch et al. (2022)'s comparative analysis of CNN and transfer learning models highlights the differential performance in detecting alligator cracks and longitudinal cracks. The paper's call for expanding longitudinal crack datasets underscores the significance of robust training data. Lee and Huh (2022)'s development of a mobile mapping system (MMS) for capturing real-time RGB and IR images of asphalt pavement surfaces showcases a fusion of sensor technology and deep learning. The paper's focus on diverse surface types and more expansive image data adds depth to the discussion of real-world applications. Lu et al. (2022)'s multi-scale crack detection network (MSCNet) with texture enhancement and feature aggregation demonstrates precision and recall rates. The paper's commitment to improving crack detection performance and inference speed aligns with the quest for accurate and efficient methodologies. Kim et al. (2022)'s Deep Bridge Crack Classification (DBCC)-Net introduces a novel approach with implications for improving Miou.

The study's recognition of limitations involving reasoning time and research data availability encourages ongoing exploration and validation.

As we conclude our journey through these papers, we embrace the diverse methodologies, insights, and advancements presented. From deep learning architectures and transfer learning to novel fusion techniques and real-world applications, this discussion underscores the multidimensional nature of crack detection research. As the field continues to evolve, these papers collectively provide a foundation for future exploration and innovation, inspiring researchers and practitioners to address challenges, bridge gaps, and strive for accurate and efficient crack detection solutions. To summarize this journey, we can say that the crack detection methodology described in the paper employs a diverse range of tools and techniques, primarily centered around deep learning algorithms and associated frameworks. These tools include.

1. Deep Learning Frameworks: The study utilizes various deep learning frameworks such as YOLO (You Only Look Once), UNet, ResNet (Residual Neural Network), Rsef, and others. These frameworks serve as the backbone for developing and training crack detection models, leveraging their capabilities in feature extraction, classification, and segmentation tasks.
2. Metrics and Evaluation Tools: To assess the performance of crack detection algorithms, the study employs a variety of metrics such as accuracy, precision, recall, F1 score, IoU (Intersection over Union), mAP (Mean Average Precision), ROC (Receiver Operating Characteristic) curve, PR (Precision-Recall) curve, Dice coefficient, and MCC (Matthews Correlation Coefficient). These metrics provide insights into the accuracy, robustness, and efficiency of the models.
3. Custom Model Implementations: The paper describes the development and implementation of custom models such as Custom YOLOv7, AttentionCrack, Rsef-Edge, Stacking Ensemble Model, Automated YOLO v4, CrackNet, CrackClassification, Mask R-CNN, EfficientNet with Residual U-Net, and others. These custom models incorporate novel architectures, attention mechanisms, and ensemble learning techniques to enhance crack detection accuracy and efficiency.
4. Data Processing and Annotation Tools: In addition to deep learning frameworks, the study may utilize various data processing and annotation tools for preprocessing raw data, labeling crack instances, and augmenting datasets. These tools ensure the quality and diversity of the training data, contributing to the robustness of the crack detection models.
5. Model Training and Optimization Tools: Model training and optimization are critical components of the crack detection methodology, requiring tools for hyperparameter tuning, optimization algorithms, and training pipelines. These tools help fine-tune model parameters, improve convergence speed, and enhance overall performance.

In summary, the crack detection methodology outlined in the paper leverages a combination of deep learning frameworks, evaluation metrics, custom model implementations, data processing tools, and model training techniques to achieve accurate and efficient crack detection results. These tools collectively enable researchers to develop, evaluate, and optimize crack detection algorithms for various real-world applications.

4 Conclusion

In this comprehensive survey paper, we embarked on a journey through the landscape of crack detection methodologies, datasets, algorithms, and results. The evolution of crack detection technologies has been remarkable, driven by the integration of cutting-edge machine learning and deep learning techniques. Our exploration revealed a diverse array of strategies, methodologies, and advancements that collectively contribute to the enhancement of structural health assessment.

As we traversed through the realms of crack detection, it became evident that traditional approaches have taken a backseat in favor of innovative and state-of-the-art methods. A prevailing trend among the surveyed articles was the utilization of contemporary variants of algorithms, such as the latest versions of YOLO, UNet, ResNet, Rsef, and more. This signifies a dynamic shift towards harnessing the full potential of modern techniques to address the complex challenges of crack detection.

A striking observation in this survey was the prevalent use of multiple algorithms within a single study. Many researchers adopted a holistic approach by combining various algorithms, with one focused on crack detection and another dedicated to segmentation. This synergy enables a more comprehensive analysis, leveraging the strengths of different methods to achieve more accurate and robust results.

Furthermore, several papers demonstrated ingenuity by devising hybrid architectures that amalgamate basic and expert models. This innovative approach capitalizes on the strengths of each model type, potentially yielding enhanced performance and adaptability in crack detection scenarios.

Authors across various papers showcased a penchant for pioneering methods and technologies in both data collection and algorithm development. This inventive spirit has led to the construction of novel datasets, precise annotations, and ingenious models tailored to the intricacies of real-world crack detection challenges.

As we peer into the future, the survey highlights the potential for further exploration and innovation. Emerging technologies like Vision Transformers (ViT) hold promise in the realm of crack detection, offering new avenues for enhancing model performance and adaptability. The integration of ViT and other groundbreaking algorithms presents exciting opportunities for researchers to push the boundaries of crack detection capabilities.

In conclusion, the amalgamation of advanced algorithms, diverse datasets, and pioneering methodologies showcased in this survey underscores the dynamic nature of crack detection research. The journey through the diverse facets of this field not only offers a deeper understanding of its current state but also inspires new horizons of exploration. As the cracks in our built environment continue to challenge us, this survey paper serves as a roadmap for researchers and practitioners, guiding them

towards the next era of innovation and excellence in crack detection technologies.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

HK: Conceptualization, Investigation, Writing–original draft, Writing–review and editing. RA: Conceptualization, Investigation, Supervision, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2024.1321634/full#supplementary-material

References

Ali, L., Jassmi, H. A., Khan, W., and Alnajjar, F. (2022). Crack45k: integration of vision transformer with tubularity flow field (tuff) and sliding-window approach for crack-segmentation in pavement structures. *Buildings* 13 (1), 55. doi:10.3390/buildings13010055

Ashraf, A., Sophian, A., Shafie, A. A., Gunawan, T. S., Ismail, N. N., and Bawono, A. A. (2023). Efficient pavement crack detection and classification using custom yolov7 model. *Indonesian J. Electr. Eng. Inf. (IJEEI)* 11 (1), 119–132. doi:10.52549/ijeei.v11i1.4362

Bai, Y., Zha, B., Sezen, H., and Yilmaz, A. (2023). Engineering deep learning methods on automatic detection of damage in infrastructure due to extreme events. *Struct. Health Monit.* 22 (1), 338–352. doi:10.1177/14759217221083649

BaniMustafa, A., AbdelHalim, R., Bulkrock, O., and Al-Hmouz, A. (2023). Deep learning for assessing severity of cracks in concrete structures. *Int. J. Comput. Commun. Control* 18 (1). doi:10.15837/ijccc.2023.1.4977

Chen, J., Yuan, Y., Lang, H., Ding, S., and Lu, J. J. (2022). The improvement of automated crack segmentation on concrete pavement with graph network. *J. Adv. Transp.* 2022, 1–10. doi:10.1155/2022/2238095

Chu, H.-H., Saeed, M. R., Rashid, J., Mehmood, M. T., Ahmad, I., Iqbal, R. S., et al. (2023). Deep learning method to detect the road cracks and potholes for smart cities. *Cmc-Computers Mater. Continua* 75 (1), 1863–1881. doi:10.32604/cmc.2023.035287

de León, G., Fiorentini, N., Leandri, P., and Losa, M. (2023). A new region-based minimal path selection algorithm for crack detection and ground truth labeling exploiting gabor filters. *Remote Sens.* 15 (11), 2722. doi:10.3390/rs15112722

Deng, L., Zhang, A., Guo, J., and Liu, Y. (2023). An integrated method for road crack segmentation and surface feature quantification under complex backgrounds. *Remote Sens.* 15 (6), 1530. doi:10.3390/rs15061530

Dorafshan, S., Thomas, R. J., and Maguire, M. (2018). Sdnet2018: an annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. *Data Brief* 21, 1664–1668. doi:10.1016/j.dib.2018.11.015

Elghaish, F., Talebi, S., Abdellatef, E., Matarneh, S. T., Hosseini, M. R., Wu, S., et al. (2022). Developing a new deep learning cnn model to detect and classify highway cracks. *J. Eng. Des. Technol.* 20 (4), 993–1014. doi:10.1108/jedt-04-2021-0192

Gharehbaghi, V., Noroozinejad Farsangi, E., Yang, T., Noori, M., and Kontoni, D.-P. N. (2022). A novel computer-vision approach assisted by 2d-wavelet transform and locality sensitive discriminant analysis for concrete crack detection. *Sensors* 22 (22), 8986. doi:10.3390/s22228986

Golding, V. P., Gharineiat, Z., Munawar, H. S., and Ullah, F. (2022). Crack detection in concrete structures using deep learning. *Sustainability* 14 (13), 8117. doi:10.3390/su14138117

Gooda, S. K., Chinthamu, N., Selvan, S. T., Rajakumareswaran, V., and Brindha, G. (2023). Automatic detection of road cracks using efficientnet with residual u-net-based segmentation and yolov5-based detection. *Int. J. Recent Innov. Trends Comput. Commun.* 11, 4. doi:10.17762/ijritcc.v11i4s.6310

Guo, Y., Shen, X., Linke, J., Wang, Z., and Barati, K. (2023). Quantification of structural defects using pixel level spatial information from photogrammetry. *Sensors* 23 (13), 5878. doi:10.3390/s23135878

Ha, J., Kim, D., and Kim, M. (2022). Assessing severity of road cracks using deep learning-based segmentation and detection. *J. Supercomput.* 78 (16), 17721–17735. doi:10.1007/s11227-022-04560-x

Hammouch, W., Chouiekh, C., Khaissidi, G., and Mrabti, M. (2022). Crack detection and classification in moroccan pavement using convolutional neural network. *Infrastructures* 7 (11), 152. doi:10.3390/infrastructures7110152

Inam, H., Islam, N. U., Akram, M. U., and Ullah, F. (2023). Smart and automated infrastructure management: a deep learning approach for crack detection in bridge images. *Sustainability* 15 (3), 1866. doi:10.3390/su15031866

Inácio, D., Oliveira, H., Oliveira, P., and Correia, P. (2023). A low-cost deep learning system to characterize asphalt surface deterioration. *Remote Sens.* 15 (6), 1701. doi:10.3390/rs15061701

Islam, M. M., Hossain, M. B., Akhtar, M. N., Moni, M. A., and Hasan, K. F. (2022). Cnn based on transfer learning models using data augmentation and transformation for detection of concrete crack. *Algorithms* 15 (8), 287. doi:10.3390/a15080287

Jayaraju, P., Somasundaram, K., Suprakash, A. S., and Muthusamy, S. (2022). A deep learning-image based approach for detecting cracks in buildings. *Trait. Du. Signal* 39 (4), 1429–1434. doi:10.18280/ts.390437

Ji, H., Kim, J., Hwang, S., and Park, E. (2022). Automated crack detection via semantic segmentation approaches using advanced u-net architecture. *Intelligent Automation Soft Comput.* 34 (1), 593–607. doi:10.32604/iasc.2022.024405

Jing, P., Yu, H., Hua, Z., Xie, S., and Song, C. (2022). Road crack detection using deep neural network based on attention mechanism and residual structure. *IEEE Access* 11, 919–929. doi:10.1109/access.2022.3233072

Jiya, E., Anwar, N., and Abdullah, M. (2016). Detection of cracks in concrete structure using microwave imaging technique. *Int. J. Microw. Sci. Technol.* 2016, 1–6. doi:10.1155/2016/3195716

Kang, D. H., and Cha, Y.-J. (2022). Efficient attention-based deep encoder and decoder for automatic crack segmentation. *Struct. Health Monit.* 21 (5), 2190–2205. doi:10.1177/14759217211053776

Kao, S.-P., Chang, Y.-C., and Wang, F.-L. (2023). Combining the yolov4 deep learning model with uav imagery processing technology in the extraction and quantization of cracks in bridges. *Sensors* 23 (5), 2572. doi:10.3390/s23052572

Kapadia, H. K., Patel, P. V., and Patel, J. B. (2023). Convolutional neural network based improved crack detection in concrete cubes. *Int. J. Comput. Digital Syst.* 13 (1), 341–352. doi:10.12785/ijcds/130127

Kim, B., Natarajan, Y., Munisamy, S. D., Rajendran, A., Sri Preethaa, K., Lee, D.-E., et al. (2022). Deep learning activation layer-based wall quality recognition using conv2d resnet exponential transfer learning model. *Mathematics* 10 (23), 4602. doi:10.3390/math10234602

Kim, J.-Y., Park, M.-W., Huynh, N. T., Shim, C., and Park, J.-W. (2023a). Detection and length measurement of cracks captured in low definitions using convolutional neural networks. *Sensors* 23 (8), 3990. doi:10.3390/s23083990

Kim, Y., Yi, S., Ahn, H., and Hong, C.-H. (2023b). Accurate crack detection based on distributed deep learning for iot environment. *Sensors* 23 (2), 858. doi:10.3390/s23020858

Kolappan Geetha, G., Yang, H.-J., and Sim, S.-H. (2023). Fast detection of missing thin propagating cracks during deep-learning- based concrete crack/non-crack classification. *Sensors* 23 (3), 1419. doi:10.3390/s23031419

Kou, L., Sysyn, M., Fischer, S., Liu, J., and Nabochenko, O. (2022). Optical rail surface crack detection method based on semantic segmentation replacement for magnetic particle inspection. *Sensors* 22 (21), 8214. doi:10.3390/s22218214

Konig, J., Jenkins, M. D., Mannion, M., Barrie, P., and Morison, G. (2021). Optimized deep encoder-decoder methods for crack segmentation. *Digit. Signal Process.* 108, 102907. doi:10.1016/j.dsp.2020.102907

Kun, J., Zhenhai, Z., Jiale, Y., and Jianwu, D. (2022). A deep learning-based method for pixel-level crack detection on concrete bridges. *IET Image Process.* 16 (10), 2609–2622. doi:10.1049/ipr2.12512

Lee, H., and Yoo, J. (2023). Fast attention cnn for fine-grained crack segmentation. *Sensors* 23 (4), 2244. doi:10.3390/s23042244

Lee, J., and Huh, Y. (2022). Multi-sensorial image dataset collected from mobile mapping system for asphalt pavement management. *Sensors Mater.* 34 (7), 2615–2624. doi:10.18494/sam3731

Lee, S., Jeong, M., Cho, C.-S., Park, J., and Kwon, S. (2022). Deep learning-based pc member crack detection and quality inspection support technology for the precise construction of osc projects. *Appl. Sci.* 12 (19), 9810. doi:10.3390/app12199810

Lee, T., Kim, J.-H., Lee, S.-J., Ryu, S.-K., and Joo, B.-C. (2023). Improvement of concrete crack segmentation performance using stacking ensemble learning. *Appl. Sci.* 13 (4), 2367. doi:10.3390/app13042367

Li, G., Liu, T., Fang, Z., Shen, Q., and Ali, J. (2022a). Automatic bridge crack detection using boundary refinement based on real-time segmentation network. *Struct. Control Health Monit.* 29 (9), e2991. doi:10.1002/stc.2991

Li, J., Lu, X., Zhang, P., and Li, Q. (2023a). Intelligent detection method for concrete dam surface cracks based on two-stage transfer learning. *Water* 15 (11), 2082. doi:10.3390/w15112082

Li, J., Tian, Y., Chen, J., and Wang, H. (2023b). Rock crack recognition technology based on deep learning. *Sensors* 23 (12), 5421. doi:10.3390/s23125421

Li, L., Fang, B., and Zhu, J. (2022b). Performance analysis of the yolov4 algorithm for pavement damage image detection with different embedding positions of cbam modules. *Appl. Sci.* 12 (19), 10180. doi:10.3390/app121910180

Li, P., Xia, H., Zhou, B., Yan, F., and Guo, R. (2022c). A method to improve the accuracy of pavement crack identification by combining a semantic segmentation and edge detection model. *Appl. Sci.* 12 (9), 4714. doi:10.3390/app12094714

Liu, H., Yang, C., Li, A., Huang, S., Feng, X., Ruan, Z., et al. (2022). Deep domain adaptation for pavement crack detection. *IEEE Trans. Intelligent Transp. Syst.* 24 (2), 1–13. doi:10.1109/tits.2022.3225212

Liu, X., Hong, Z., Shi, W., and Guo, X. (2023). Image-processing-based subway tunnel crack detection system. *Sensors* 23 (13), 6070. doi:10.3390/s23136070

Loverdos, D., and Sarhosis, V. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. *Automation Constr.* 140, 104389. doi:10.1016/j.autcon.2022.104389

Lu, G., He, X., Wang, Q., Shao, F., Wang, J., and Zhao, X. M. (2022). MSCNet: a framework with a texture enhancement mechanism and feature aggregation for crack detection. *IEEE Access* 10, 26127–26139. doi:10.1109/access.2022.3156606

Lv, Z., Cheng, C., and Lv, H. (2023). Automatic identification of pavement cracks in public roads using an optimized deep convolutional neural network model. *Philosophical Trans. R. Soc. A* 381 (2254), 20220169. doi:10.1098/rsta.2022.0169

Ma, D., Fang, H., Wang, N., Xue, B., Dong, J., and Wang, F. (2022a). A real-time crack detection algorithm for pavement based on cnn with multiple feature layers. *Road Mater. Pavement Des.* 23 (9), 2115–2131. doi:10.1080/14680629.2021.1925578

Ma, J., Yan, W., Liu, G., Xing, S., Niu, S., and Wei, T. (2022b). Complex texture contour feature extraction of cracks in timber structures of ancient architecture based on yolo algorithm. *Adv. Civ. Eng.* 2022, 1–13. doi:10.1155/2022/7879302

Maslan, J., and Cicmanec, L. (2023). A system for the automatic detection and evaluation of the runway surface cracks obtained by unmanned aerial vehicle imagery using deep convolutional neural networks. *Appl. Sci.* 13 (10), 6000. doi:10.3390/app13106000

Mo, D.-H., Wu, Y.-C., and Lin, C.-S. (2022). The dynamic image analysis of retaining wall crack detection and gap hazard evaluation method with deep learning. *Appl. Sci.* 12 (18), 9289. doi:10.3390/app12189289

Mohammed, M. A., Han, Z., Li, Y., Al-Huda, Z., Li, C., and Wang, W. (2022). End-to-end semi-supervised deep learning model for surface crack detection of infrastructures. *Front. Mater.* 9, 1058407. doi:10.3389/fmats.2022.1058407

Munawar, H., Ullah, F., Heravi, A., Thaheem, M., and Maqsoom, A. (2022a). Inspecting buildings using drones and computer vision: a machine learning approach to detect cracks and damages. *Drones* 6, 5. doi:10.3390/drones6010005

Munawar, H. S., Ullah, F., Shahzad, D., Heravi, A., Qayyum, S., and Akram, J. (2022b). Civil infrastructure damage and corrosion detection: an application of machine learning. *Buildings* 12 (2), 156. doi:10.3390/buildings12020156

Ngo, L., Xuan, C. L., Luong, H. M., Thanh, B. N., and Ngoc, D. B. (2023). Designing image processing tools for testing concrete bridges by a drone based on deep learning. *J. Inf. Telecommun.* 7 (2), 227–240. doi:10.1080/24751839.2023.2186624

Nomura, Y., Inoue, M., and Furuta, H. (2022). Evaluation of crack propagation in concrete bridges from vehicle-mounted camera images using deep learning and image processing. *Front. Built Environ.* 8, 972796. doi:10.3389/fbuil.2022.972796

O¨ zgenel, C. F., and Sorguc, A. G. (2018). "Performance comparison of pretrained convolutional neural networks on crack detection in buildings," in Isarc. proceedings of the international symposium on automation and robotics in construction, Berlin, Germany, 2018, (IAARC Publications), 1–8.

Panta, M., Hoque, M. T., Abdelguerfi, M., and Flanagin, M. C. (2023). Iterlunet: deep learning architecture for pixel-wise crack detection in levee systems. *IEEE Access* 11, 12249–12262. doi:10.1109/access.2023.3241877

Paramanandham, N., Koppad, D., and Anbalagan, S. (2022). Vision based crack detection in concrete structures using cutting-edge deep learning techniques. *Trait. Du. Signal* 39 (2), 485–492. doi:10.18280/ts.390210

Paramanandham, N., Rajendiran, K., Poovathy J, F. G., Premanand, Y. S., Mallichetty, S. R., and Kumar, P. (2023). Pixel intensity resemblance measurement and deep learning based computer vision model for crack detection and analysis. *Sensors* 23 (6), 2954. doi:10.3390/s23062954

Philip, R. E., Andrushia, A. D., Nammalvar, A., Gurupatham, B. G. A., and Roy, K. (2023). A comparative study on crack detection in concrete walls using transfer learning techniques. *J. Compos. Sci.* 7 (4), 169. doi:10.3390/jcs7040169

Popli, R., Kansal, I., Verma, J., Khullar, V., Kumar, R., and Sharma, A. (2023). Road: robotics-assisted onsite data collection and deep learning enabled robotic vision system for identification of cracks on diverse surfaces. *Sustainability* 15 (12), 9314. doi:10.3390/su15129314

Pu, R., Ren, G., Li, H., Jiang, W., Zhang, J., and Qin, H. (2022). Autonomous concrete crack semantic segmentation using deep fully convolutional encoder–decoder network in concrete structures inspection. *Buildings* 12 (11), 2019. doi:10.3390/buildings12112019

Qayyum, W., Ehtisham, R., Bahrami, A., Camp, C., Mir, J., and Ahmad, A. (2023). Assessment of convolutional neural network pre-trained models for detection and orientation of cracks. *Materials* 16 (2), 826. doi:10.3390/ma16020826

Quqa, S., Martakis, P., Movsessian, A., Pai, S., Reuland, Y., and Chatzi, E. (2022). Two-step approach for fatigue crack detection in steel bridges using convolutional neural networks. *J. Civ. Struct. Health Monit.* 12 (1), 127–140. doi:10.1007/s13349-021-00537-1

Ren, J., Zhao, G., Ma, Y., Zhao, D., Liu, T., and Yan, J. (2022). Automatic pavement crack detection fusing attention mechanism. *Electronics* 11 (21), 3622. doi:10.3390/electronics11213622

Shim, S., Kim, J., Cho, G.-C., and Lee, S.-W. (2023). Stereo-vision-based 3d concrete crack detection using adversarial learning with balanced ensemble discriminator networks. *Struct. Health Monit.* 22 (2), 1353–1375. doi:10.1177/14759217221097868

Siriborvornratanakul, T. (2022). Downstream semantic segmentation model for low-level surface crack detection. *Adv. Multimedia* 2022, 1–12. doi:10.1155/2022/3712289

Sun, Z., Zhai, J., Pei, L., Li, W., and Zhao, K. (2023). Automatic pavement crack detection transformer based on convolutional and sequential feature fusion. *Sensors* 23 (7), 3772. doi:10.3390/s23073772

Tan, H., and Dong, S. (2023). Pixel-level concrete crack segmentation using pyramidal residual network with omni-dimensional dynamic convolution. *Processes* 11 (2), 546. doi:10.3390/pr11020546

Tang, W., Huang, S., Zhao, Q., Li, R., and Huangfu, L. (2021). An iteratively optimized patch label inference network for automatic pavement distress detection. *IEEE Trans. Intelligent Transp. Syst.* 23, 8652–8661. doi:10.1109/tits.2021.3084809

Tse, K.-W., Pi, R., Sun, Y., Wen, C.-Y., and Feng, Y. (2023). A novel real-time autonomous crack inspection system based on unmanned aerial vehicles. *Sensors* 23 (7), 3418. doi:10.3390/s23073418

Wan, C., Xiong, X., Wen, B., Gao, S., Fang, D., Yang, C., et al. (2022). Crack detection for concrete bridges with imaged based deep learning. *Sci. Prog.* 105 (4), 003685042211284. doi:10.1177/00368504221128487

Wang, J.-J., Liu, Y.-F., Nie, X., and Mo, Y. (2022a). Deep convolutional neural networks for semantic segmentation of cracks. *Struct. Control Health Monit.* 29 (1), e2850. doi:10.1002/stc.2850

Wang, L. (2023). Automatic detection of concrete cracks from images using adam-squeezenet deep learning model. *Frat. Ed. Integritá Strutt.* 17 (65), 289–299. doi:10.3221/igf-esis.65.19

Wang, Y., Wang, J., Wang, C., Wen, X., Yan, C., Guo, Y., et al. (2022b). Ma-xnet: mobile-attention x-network for crack detection. *Appl. Sci.* 12 (21), 11240. doi:10.3390/app122111240

Wibowo, A. P., Adha, A., Kurniawan, I. F., and Laory, I. (2022). Wall crack multiclass classification: expertise-based dataset construction and learning algorithms performance comparison. *Buildings* 12 (12), 2135. doi:10.3390/buildings12122135

Wu, D., Zhang, H., and Yang, Y. (2022). Deep learning-based crack monitoring for ultra-high performance concrete (uhpc). *J. Adv. Transp.* 2022, 1–10. doi:10.1155/2022/4117957

Xu, X., Zhao, M., Shi, P., Ren, R., He, X., Wei, X., et al. (2022). Crack detection and comparison study based on faster r-cnn and mask r-cnn. *Sensors* 22 (3), 1215. doi:10.3390/s22031215

Yadav, D. P., Kishore, K., Gaur, A., Kumar, A., Singh, K. U., Singh, T., et al. (2022). A novel multi-scale feature fusion-based 3scnet for building crack detection. *Sustainability* 14 (23), 16179. doi:10.3390/su142316179

Yang, N., Li, Y., and Ma, R. (2022). An efficient method for detecting asphalt pavement cracks and sealed cracks based on a deep data-driven model. *Appl. Sci.* 12 (19), 10089. doi:10.3390/app121910089

Yang, Y., Xu, W., Zhu, Y., Su, L., and Zhang, G. (2023). A novel detection method for pavement crack with encoder-decoder architecture. *CMES-Computer Model. Eng. Sci.* 137 (1), 761–773. doi:10.32604/cmes.2023.027010

Yong, P., and Wang, N. (2022). Riianet: a real-time segmentation network integrated with multi-type features of different depths for pavement cracks. *Appl. Sci.* 12 (14), 7066. doi:10.3390/app12147066

Yu, G., and Zhou, X. (2023). An improved yolov5 crack detection method combined with a bottleneck transformer. *Mathematics* 11 (10), 2377. doi:10.3390/math11102377

Yu, J., Wu, C., Li, Y., and Zhang, Y. (2022). Intelligent identification of coal crack in ct images based on deep learning. *Comput. Intell. Neurosci.* 2022. doi:10.1155/2022/7092436

Yuan, B., Sun, Z., Pei, L., Li, W., Ding, M., and Hao, X. (2022). Super-resolution reconstruction method of pavement crack images based on an improved generative adversarial network. *Sensors* 22 (23), 9092. doi:10.3390/s22239092

Zhang, C., Chen, Y., Tang, L., Chu, X., and Li, C. (2023a). Ctcd-net: a cross-layer transmission network for tiny road crack detection. *Remote Sens.* 15 (8), 2185. doi:10.3390/rs15082185

Zhang, J., Qian, S., and Tan, C. (2023b). Automated bridge crack detection method based on lightweight vision models. *Complex and Intelligent Syst.* 9 (2), 1639–1652. doi:10.1007/s40747-022-00876-6

Zhang, L., Yang, F., Zhang, Y. D., and Zhu, Y. J. (2016). "Road crack detection using deep convolutional neural network," in 2016 IEEE international conference on image processing (ICIP) (IEEE), 3708–3712.

Zhang, Y., Zhang, Z., Zhao, W., and Li, Q. (2022). Crack segmentation on earthen heritage site surfaces. *Appl. Sci.* 12 (24), 12830. doi:10.3390/app122412830

Zhao, F., Chao, Y., and Li, L. (2023a). A crack segmentation model combining morphological network and multiple loss mechanism. *Sensors* 23 (3), 1127. doi:10.3390/s23031127

Zhao, M., Shi, P., Xu, X., Xu, X., Liu, W., and Yang, H. (2022). Improving the accuracy of an r-cnn-based crack identification system using different preprocessing algorithms. *Sensors* 22 (18), 7089. doi:10.3390/s22187089

Zhao, Y., Zhou, L., Wang, X., Wang, F., and Shi, G. (2023b). Highway crack detection and classification using uav remote sensing images based on cracknet and crackclassification. *Appl. Sci.* 13 (12), 7269. doi:10.3390/app13127269

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 06 August 2024
DOI 10.3389/fenvs.2024.1416765

OPEN ACCESS

EDITED BY
Sayali Apte,
Symbiosis International University, India

REVIEWED BY
F. E. Jalal,
Shenzhen University, China
Daniela Varrica,
University of Palermo, Italy

*CORRESPONDENCE
Kai Zhang,
hugo88315@163.com
Qiong Pan,
panqion@nwsuaf.edu.cn

RECEIVED 13 April 2024
ACCEPTED 22 July 2024
PUBLISHED 06 August 2024

CITATION
Wang X, Zhang K, Han P, Wang M, Li X, Zhang Y and Pan Q (2024) Application of gene expression programing in predicting the concentration of $PM_{2.5}$ and PM_{10} in Xi'an, China: a preliminary study.
Front. Environ. Sci. 12:1416765.
doi: 10.3389/fenvs.2024.1416765

Application of gene expression programing in predicting the concentration of $PM_{2.5}$ and PM_{10} in Xi'an, China: a preliminary study

Xu Wang[1], Kai Zhang[2]*, Peishan Han[3], Meijia Wang[4], Xianjun Li[5], Yaqiong Zhang[6] and Qiong Pan[7]*

[1]School of Software, Shanxi Agricultural University, Taiyuan, China, [2]Chongqing Chang'an Industrial Co., Ltd., Chongqing, China, [3]Passenger Transport Third Branch, Shenzhen Metro Operation Group Co., Ltd., Shenzhen, China, [4]School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an Weiyang University Park, Xi'an, China, [5]Meteorological Bureau of Yangling, Yangling, Shaanxi, China, [6]Endoscopy Center, Minhang Hospital, Fudan University, Shanghai, China, [7]School of Science, Northwestern A&F University, Yangling, China

Introduction: Traditional statistical methods cannot find quantitative relationship from environmental data.

Methods: We selected gene expression programming (GEP) to study the relationship between pollutant gas and $PM_{2.5}$ (PM_{10}). They were used to construct the relationship between pollutant gas and $PM_{2.5}$ (PM_{10}) with environmental monitoring data of Xi'an, China. GEP could construct a formula to express the relationship between pollutant gas and $PM_{2.5}$ (PM_{10}), which is more explainable. Back Propagation neural networks (BPNN) was used as the baseline method. Relevant data from January 1st 2021 to April 26th 2021 were used to train and validate the performance of the models from GEP and BPNN.

Results: After the models of GEP and BPNN constructed, coefficient of determination and RMSE (Root Mean Squared Error) are used to evaluate the fitting degree and measure the effect power of pollutant gas on $PM_{2.5}$ (PM_{10}). GEP achieved RMSE of [8.7365–14.6438] for $PM_{2.5}$; RMSE of [13.2739–45.8769] for PM_{10}, and BP neural networks achieved average RMSE of [13.8741–34.7682] for $PM_{2.5}$; RMSE of [29.7327–52.8653] for PM_{10}. Additionally, experimental results show that the influence power of pollutant gas on $PM_{2.5}$ (PM_{10}) situates between −0.0704 and 0.6359 (between −0.3231 and 0.2242), and the formulas are obtained with GEP so that further analysis become possible. Then linear regression was employed to study which pollutant gas is more relevant to $PM_{2.5}$ (PM_{10}), the result demonstrates CO (SO_2, NO_2) are more related to $PM_{2.5}$ (PM_{10}).

Discussion: The formulas produced by GEP can also provide a direct relationship between pollutant gas and $PM_{2.5}$ (PM_{10}). Besides, GEP could model the trend of $PM_{2.5}$ and PM_{10} (increase and decrease). All results show that GEP can be applied smoothly in environmental modelling.

KEYWORDS

pollutant gas, $PM_{2.5}$, PM_{10}, gene expression programing, back propagation neural network, Xi'an

Introduction

PM (particulate matter) has become a dangerous threat for the health of human beings (Nel, 2005; Sun et al., 2013; Apte et al., 2015; Bossmann et al., 2016). $PM_{2.5}$ or PM_{10} are particles with a diameter less than 2.5 μm or 10 μm (Francesca Dominici et al., 2014; Pui et al., 2014) (Ostro et al., 2006; Ma et al., 2011), which have adverse health effects on respiratory health and cause more complications. The formation mechanism and process for $PM_{2.5}$ or PM_{10} are pretty complex. Major sources of $PM_{2.5}$ and PM_{10} include natural sources (plant division and spore, soil dust, sea salt, forest fire, volcano eruption and so on) and artificial sources (combustion of fuel, emission of industrial production process, and emission of transportation and so on); all these can be divided into disposable particles (particles that are emitted from the emission source directly) and secondary particles (particles that are released from the chemical reaction of emission and composition of the atmosphere). They mainly consist of water-soluble ions, particulate organic matter, and trace elements. Gautam et al. (2016) considers that NO_2, SO_2, CO, and O_3 are the main gaseous materials which can influence the concentrations of $PM_{2.5}$ and PM_{10} under certain environmental conditions, so finding the association between pollutant gases and $PM_{2.5}$ (PM_{10}) is of importance.

Because of the adverse effects caused by $PM_{2.5}$ and PM_{10} in many aspects, they are hot topics for research. Although many research studies have made plenty of achievements, the main research method is regression, time-series regression, or some existing mathematical models. In addition, the models adopted in some research studies can only produce qualitative results without direct interoperability (shown as a formula), whereas the current research adopts gene expression programing (GEP), which can effectively avoid the subjectivity of the empirical model and obtain quantitative results. We focused on modeling the relationship between $PM_{2.5}$ (PM_{10}) and pollutant gases. There are totally five types of pollutant gases, namely, SO_2, NO_2, CO, average concentration of ozone in 1 hour, and average concentration of ozone in 8 hours. We collected relevant data from Xi'an, Shaanxi province in China, which is seriously threatened by PM, and then applied GEP to complete this task. The back propagation neural network (BPNN) was used as the baseline method. Experimental results indicate that the average influence power of pollutant gases on $PM_{2.5}$ (PM_{10}) ranges from -0.0704 to 0.6359 (from -0.3231 to 0.2242), and at the same time, $PM_{2.5}$ is more seriously affected by pollutant gases than PM_{10}. Furthermore, the formulas obtained by GEP can portray the relationship and evolution law between pollutant gases and $PM_{2.5}$ (PM_{10}), and these results can be applied to predict the concentrations of $PM_{2.5}$ and PM_{10}. Furthermore, these formulas can provide more conclusions about this problem with the assistance of mathematical analysis, such as the effect of weather or season on $PM_{2.5}$ (PM_{10}), and even the above methods can be applied in this field for other perspectives. Finally, linear regression was used to study which pollutant gas more seriously influences the concentrations of $PM_{2.5}$ and PM_{10}. Experimental results show that different pollutant gases affect $PM_{2.5}$ and PM_{10} concentrations with varying degrees, especially CO and SO_2 contribute to $PM_{2.5}$ more and NO_2 is more relevant to PM_{10} overall. More data, including more abundant information, need to be employed to build a more generalized model that can help researchers control air pollutants and study the change of $PM_{2.5}$ (PM_{10}). Experimental results show that GEP can be used in environmental modeling to uncover essential laws hidden in environmental data.

Methods

GEP

Gene expression programing (GEP) was proposed by a Portuguese scholar, Candida, in 2001 on the basis of genetic algorithms (GA) and genetic programing (GP) (Ferreira, 2001a). GEP adopts a dual structure (genotype and phenotype), which retains the advantages of GA and GP, while at the same time avoiding their shortcomings. GEP has many advantages, such as a concise algorithm flow, as shown in Figure 1A, simple implementation, high precision, and exceptional performance in complex function finding problems with large amounts of data (Özcan, 2012; Mostafa and El-Masry, 2016). GEP uses computers to create a virtual creature population that consists of some chromosomes to simulate the genetic and evolution processes of creatures that can be carried out with a series of simulation genetic operations (e.g., cross-over, mutation, and selection as the fitness) during multi-generation iterations to guarantee that the virtual population evolves to the global optima. Cross-over, mutation, and selection simulate the reproduction, mutation, and natural selection processes. GEP has an ingenious individual encoding method, which is uncomplicated and makes subsequent genetic operations convenient to implement. It then applies the outstanding computing performance of computers to iteratively calculate and obtain an optimal function model. GEP has been fully applied in many fields, such as human body mechanics (Yang et al., 2016), water conservancy (Azamathulla, 2012) robotics (Wu et al., 2013), and agriculture (Yassin et al., 2016).

As the inheritance and expansion of GA and GP, GEP integrates the advantages of both and has a more powerful ability to solve problems. The GEP algorithm can be defined as a nine-meta group: $GEP = \{C, E, P_0, M, \varphi, \Gamma, \Phi, \Pi, T\}$, where C is the individual's coding means; E is the individual's fitness evaluating function; P_0 is the initial population; M is the size of the population; φ is the selection operator; Γ is the crossover operator; Φ is the point mutation operator; Π is the string mutation operator; and T is the termination condition. In GEP, each individual is also known as a chromosome, which is formed by genes that are linked together by the link operator. The gene is a linear symbol string which is composed of a head and a tail. The head involves the variables that come from the variable set (in this problem, the variable set represents the five types of pollutant gases) and the functions that come from the function set, which can be set ahead, but the tail only contains the variables that come from the variable set. The relationship between the length of the head (h) and the length of the tail (t) is expressed as $t = h^*(n-1)+1$, where n is the maximum number of parameters of each function in the function set. Each individual (chromosome) in the population could be expressed as a formula (e.g., Figure 1B shows a formula $\sqrt{(a-b)^*(c+d)}$), and the population could evolve according to the fitness evaluation function via genetic operations (Γ, Φ, Π). The genetic operations include point mutation, string mutation, and recombination. The rules of point mutation are as follows (shown in Figure 1C): the first element in the head could only be mutated to be a function in the function set, the other elements

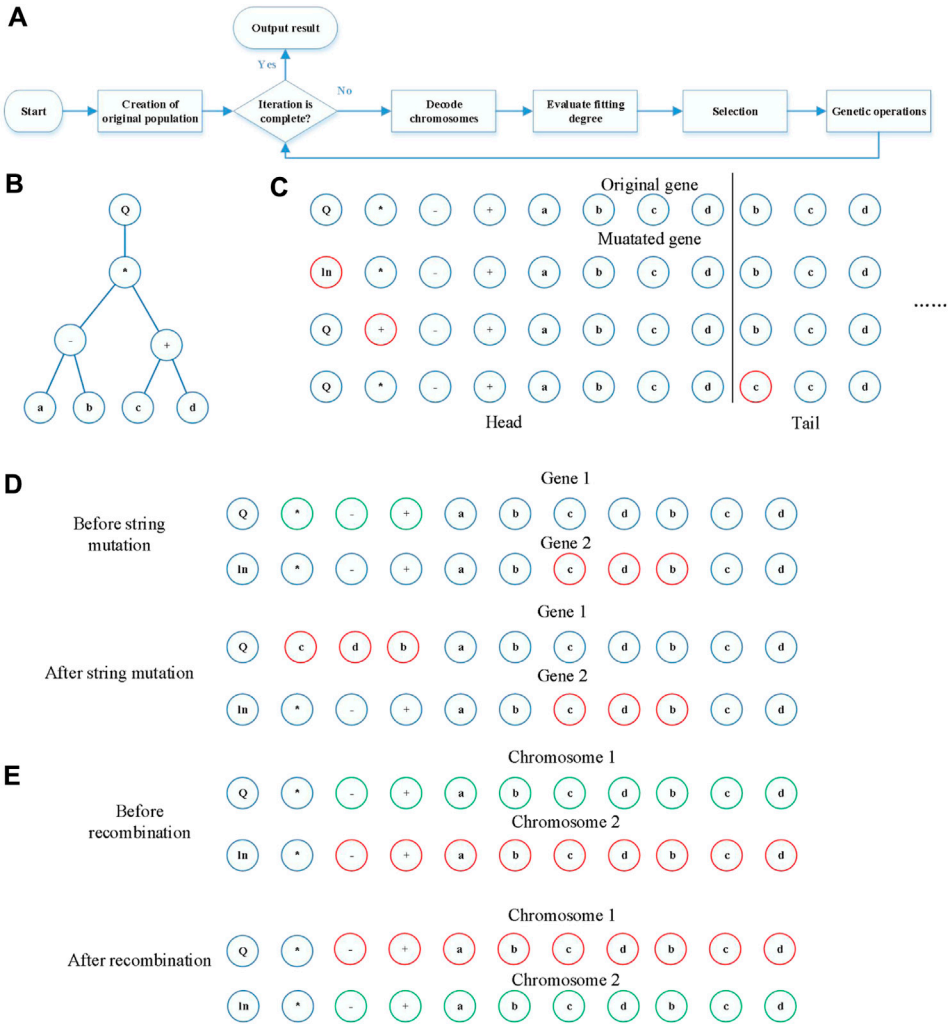


FIGURE 1
Details of GEP. **(A)** Flowchart of GEP for constructing the relationship between pollutant gases and $PM_{2.5}$ (PM_{10}). **(B)** The tree architecture of a gene. **(C)** The mutation operation: we show this operation in terms of a gene, and the green element is mutated to be the red ones. **(D)** String mutation: the green substring is mutated to be the red substring; here, we only show the insertion sequence transposition. **(E)** Recombination: the green substring and the red substring exchanged; here, we only show the single-point recombination.

in the head could be mutated to be a function in the function set or a variable in the variable set, and the elements in the tail could only be mutated to be a variable in the variable set. The rule of string mutation is as follows: a substring in a gene is replaced by a substring in another gene within a chromosome, including insertion sequence transposition, root transposition, and gene transposition (Figure 1D). The rule of recombination is as follows: the same positions of two chromosomes are exchanged, including single-point recombination, two-point recombination, and gene combination (Figure 1E). The selection operation is roulette wheel sampling, which means the chromosome showing better fitness could be selected with a higher probability and the chromosome showing worse fitness could be selected with a lower probability. This could improve the population diversity, which is good for evolutionary computation (Ferreira, 2001b). For the problem given in this study, the result from GEP is probably a function $f(x_1, x_2, x_3, x_4, x_5) = \tan(x_1) + |x_2| + \log_2(x_3/x_5) + x_4/2$ which can perfectly demonstrate the relationship between the concentration of $PM_{2.5}$ (PM_{10}) and the concentrations of five types of pollutant gases (x_1, x_2, x_3, x_4, and x_5) globally.

BP neural network

The BP neural network (Liu et al., 2015), whose structure is shown in Figure 2, is a commonly used artificial network architecture. The BP neural network makes use of fully connected neurons to form a feedforward network and then adjusts the weights of each pair of connections and the biased value of each neuron with a gradient descent algorithm that is based on the chain law of derivatives.

Given the training dataset $[x_1, d_1; ...x_i, d_i; ...; x_n, d_n]$, where x_i and d_i are the independent variable vector and dependent variable vector, respectively. There are two procedures in the training process of the BP neural network: first, the output of the

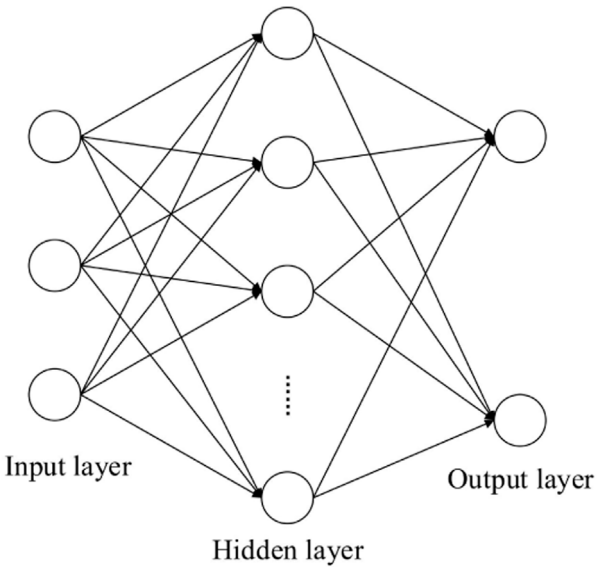


FIGURE 2
Architecture of the BP neural network.

kth neuron in the output layer is shown in Eq. 1, where w_{ij}, w'_{jk}, b_j, b'_k, and $f(\cdot)$ are the weights between the ith input neuron and the jth hidden neuron, the weight between the jth hidden neuron and the kth output neuron, the biased value of the jth hidden neuron, the biased value of the kth output neuron, and the activation function of the hidden layer, respectively; second, the weights are adjusted according to Eq. 3, which originates from the chain rule and error function e, which is shown in Eq. 2, where $\delta_k = d_k - output_k$. These two procedures are repeated until the error function converges. In this research, the input and the output of the neural network which is applied to obtain an approximate numerical regression model depicting the association between five types of pollutant gases and $PM_{2.5}$ (PM_{10}) are the concentrations of five types of pollutant gases and the concentration of $PM_{2.5}$ (PM_{10}).

$$output_k = \sum_k w'_{jk} f\left(\sum_j w_{ij} \cdot x_i + b_j\right) + b'_k, \tag{1}$$

$$e = \frac{1}{2}\sum_{i=1}^{n}(d_i - output_i)^2, \tag{2}$$

$$\Delta w'_{jk} = -\eta\frac{\partial e}{\partial w'_{jk}} = \eta(d_k - output_k) f'(net_k)$$
$$\Delta w_{ij} = -\eta\frac{\partial e}{\partial w_{ij}} = \eta\left(\sum_{k=1}^{L}\delta_k w'_{jk}\right) f'(net_j) \tag{3}$$

where η is the learning rate which is set in advance. Because the neural network can fit any nonlinear function with enough neurons, it has been widely applied in many fields, such as energy (Yu and Xu, 2014), safety (Wang L. et al., 2015; Wang Y. et al., 2015), and material science (Zhou et al., 2015).

FIGURE 3
Probable locations of five monitoring sites.

TABLE 1 Fitting degrees of GEP and BP neural network with testing data.

GEP						
Dataset	$PM_{2.5}$			PM_{10}		
	Maximum	Minimum	Mean	Maximum	Minimum	Mean
Monitoring site 1	0.8152	0.3391	0.6359	0.5071	-0.4984	0.2242
Monitoring site 2	0.6406	0.0571	0.4158	0.3453	-0.5467	0.1335
Monitoring site 3	0.5020	-0.2149	0.2091	0.2539	-0.8229	-0.0238
Monitoring site 4	0.6283	-0.1999	0.2711	0.6798	-1.2659	-0.3231
Monitoring site 5	0.6146	0.0902	0.3510	0.4878	-0.2479	0.1152
Average value of whole city	0.6918	-0.0196	0.4280	0.3769	-0.3912	0.1277
BP neural network						
Dataset	$PM_{2.5}$			PM_{10}		
	Maximum	Minimum	Mean	Maximum	Minimum	Mean
Monitoring site 1	0.4072	-0.0139	0.1979	0.3919	-0.0730	0.1989
Monitoring site 2	0.3897	-0.6140	-0.0285	0.2923	-0.2153	0.0262
Monitoring site 3	0.2128	-0.2436	0.0462	0.3928	-0.9429	-0.0925
Monitoring site 4	0.3163	-0.0504	0.0825	0.2032	-0.7365	-0.1105
Monitoring site 5	0.2827	-0.6901	-0.0704	0.4892	-2.4323	-0.2090
Average value of whole city	0.5796	-0.5143	0.1075	0.2550	-0.2479	-0.0152

Linear regression

Similar to the BP neural network, linear regression (Frank et al., 2015) is also employed to find a linear expression that portrays the relationship between independent variables and dependent variables and is expressed as Eq. 4. It adopts least squares as the objective function to minimize the error function, which is the same as the function in the BP neural network. The coefficients of each independent variable can reflect the relevance and the degree of correlation with dependent variables. As a basic data mining and intelligence information processing technique, many achievements have been made with it, such as energy (Kicsiny, 2014; Wang et al., 2016) and mechanism (Tosun et al., 2016).

$$F = a_0 + a_1{}^*x_1 + \ldots + a_n{}^*x_n, \quad (4)$$

where $x_1 - x_n$ stands for the independent variables (the concentrations of CO, SO_2, NO_2, O_3 within 1 hour, and O_3 within 8 hours); F refers to dependent variables (the concentrations of $PM_{2.5}$ or PM_{10}). In the present study, n = 5 means five types of pollutant gases.

Results

Dataset

Xi'an, which is located in northwestern China, is badly affected by PM, and we chose the samples from Xi'an as the research material. Each sample consists of the concentrations of CO, SO_2, NO_2, $PM_{2.5}$, PM_{10}, and the average concentrations of O_3 (Jerrett et al., 2009) in 1 hour or 8 hours. The dataset is collected from the environmental website of Xi'an city[1] (from 1 January 2021 to 26 April 2021), where the data are obtained from 13 monitoring sites. We selected six groups of datasets collected at five monitoring sites uniformly distributed in Xi'an and an overall average dataset of the whole city. The approximate locations of five monitoring sites are shown as Figure 3. Moreover, the incomplete samples with missing value(s) are deleted to facilitate study. For each site, three-fourth entries were used to train and one-fourth entries were randomly used to validate the model. For sites 1, 2, 3, 4, 5, and the average value of the whole city, there are 96, 104, 76, 100, 88, and 116 entries, respectively. The inputs are five types of pollutant gases (CO, SO_2, average concentration of O_3 in 8 hours, NO_2, and average concentration of O_3 in 1 hour), and the output is the concentration of $PM_{2.5}$ (PM_{10}).

Fitting degree evaluation function

In statistics, the coefficient of determination $1 - SSE/SST$ (Kumar et al., 2020) is usually used to assess the relevance degree between two groups of data, where SSE and SST is shown as Equations 5, 6

1 https://aqicn.org/city/china/xian/wentiju/cn/

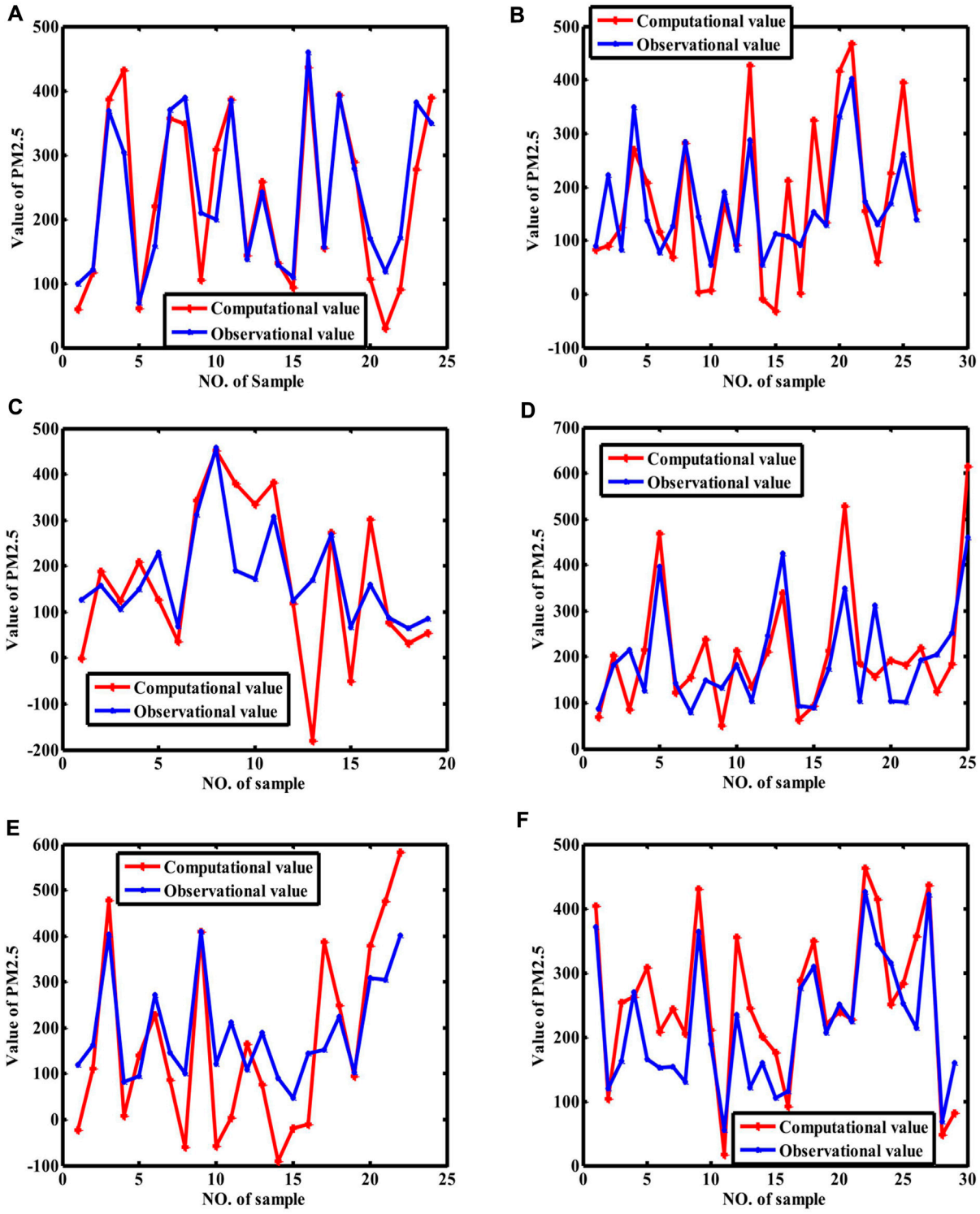


FIGURE 4
Fitting curves of $PM_{2.5}$ (GEP). (Note: **(A–F)** stand for the fitting curves obtained with datasets collected at monitoring sites 1, 2, 3, 4, 5, and average value of whole city, respectively). Because the samples are randomly selected for training and validated for repeating ten times, we only show the results with the best coefficient of determination.

$$SSE = \sum_{j=1}^{m} \left(y_j - \hat{y}_j\right)^2, \tag{5}$$

$$SST = \sum_{j=1}^{m} \left(y_j - \bar{y}\right)^2, \tag{6}$$

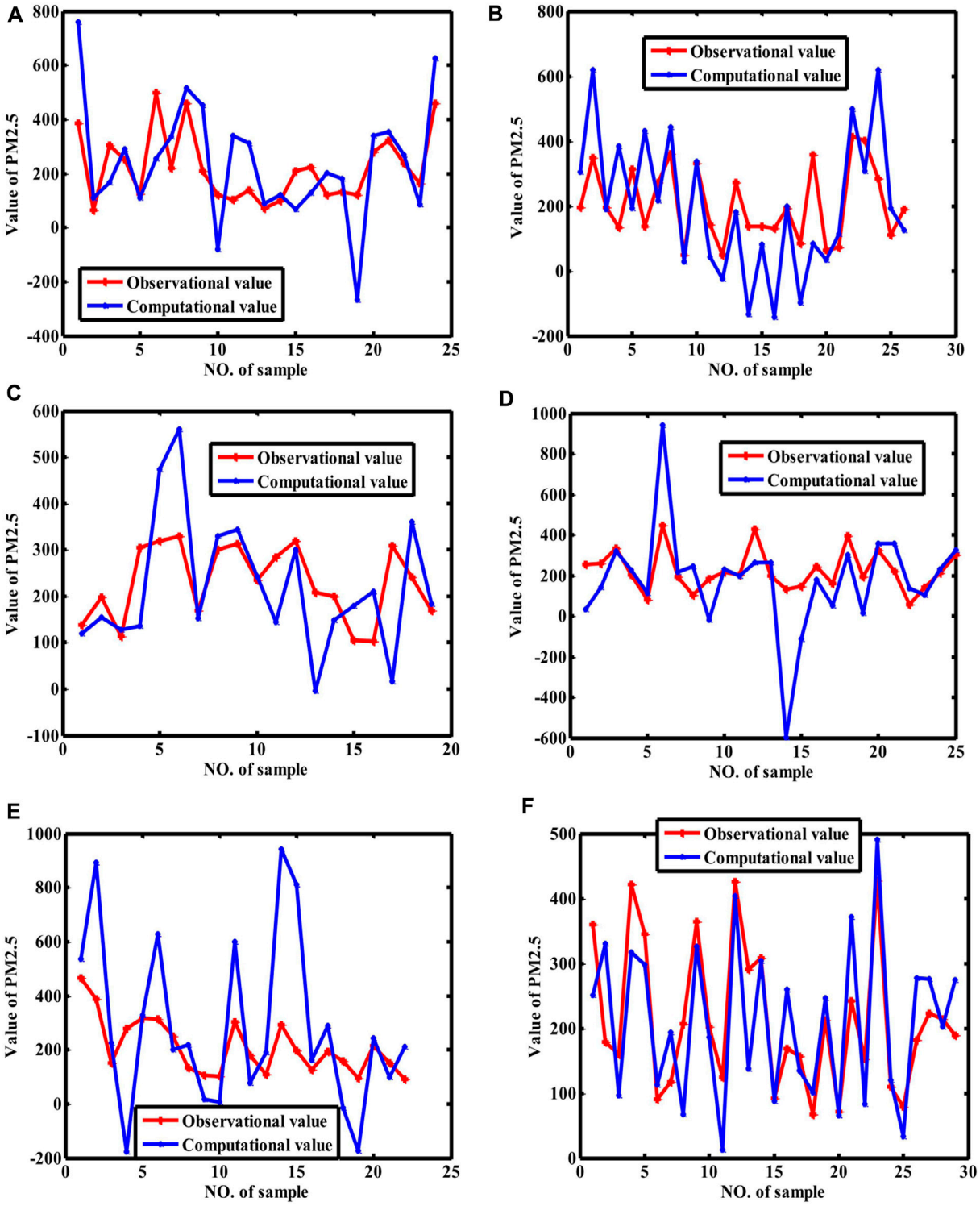


FIGURE 5
Fitting curves of $PM_{2.5}$ (BP neural network). [Note: **(A–F)** stand for the fitting curves obtained with datasets collected at monitoring sites 1, 2, 3, 4, 5, and average value of whole city, respectively]. Because the samples are randomly selected for training and validated for repeating ten times, we only show the results with the best coefficient of determination.

where y_j is the observational value of $PM_{2.5}$ (PM_{10}) and $\hat{y}_j$ is the computational value, which is computed with formulas obtained with GEP (models obtained with the BP neural network) and observation values of five types of pollutant gases. $\bar{y}$ is the average value of y. SSE is the residual sum of squares; SST is the total sum of squares of deviations. Higher values indicate higher degrees of model fitting. On the other hand, it can be used to assess the influence power of pollutant gases on $PM_{2.5}$ and PM_{10}.

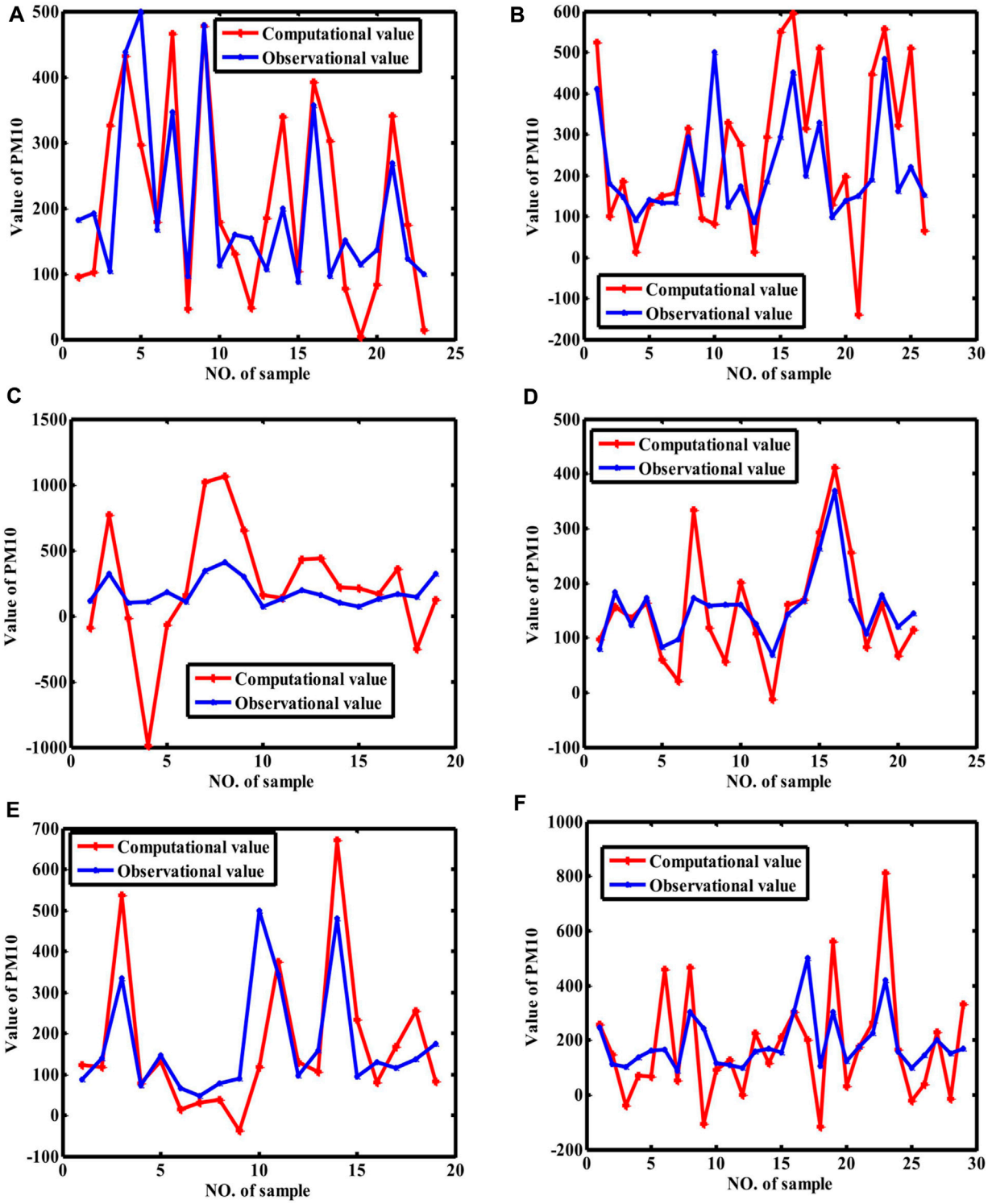


FIGURE 6
Fitting curves of PM_{10} (GEP). [Note: **(A–F)** stand for the fitting curves obtained with datasets collected at monitoring sites 1, 2, 3, 4, 5, and average value of whole city, respectively]. Because the samples are randomly selected for training and validated for repeating ten times, we only show the result with best the coefficient of determination.

Experimental settings and experimental results

First of all, the GEP (Schmidt and Lipson, 2009) and BP neural network were used to model the influence of pollutant gases on $PM_{2.5}$ and PM_{10}. All the methods were implemented using MATLAB R2016a on a personal computer with an Intel 2.80 GHz i5 processor and 8G RAM. GEP's initial step and genetic operations were randomly implemented as probabilities, and the weights and biased values of BP neural network were also

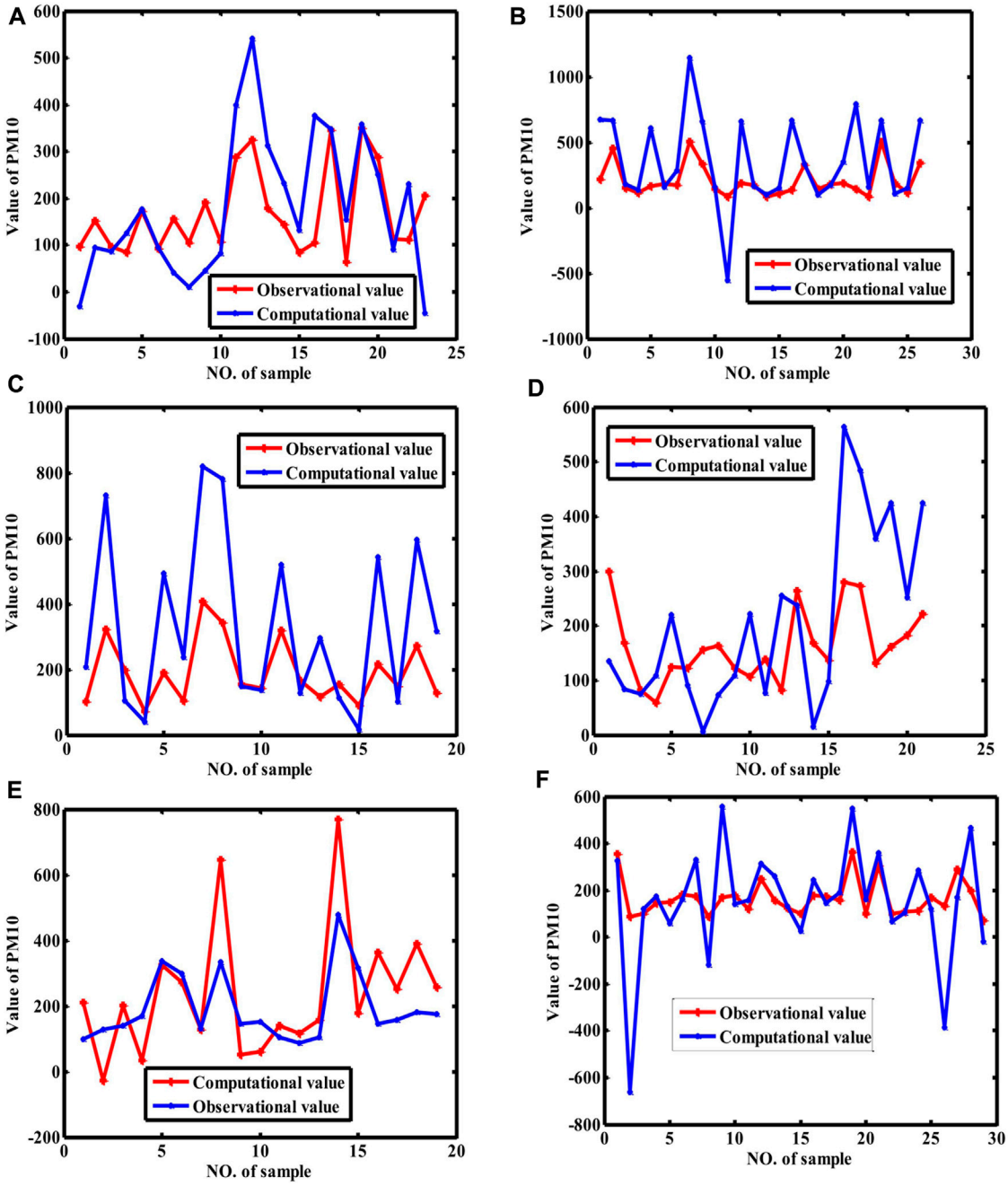


FIGURE 7
Fitting curves of PM_{10} (BP neural network). [Note: **(A–F)** stand for the fitting curves obtained with datasets collected at monitoring sites 1, 2, 3, 4, 5, and average value of whole city, respectively]. Because the samples are randomly selected for training and validated for repeating ten times, we only show the results with the best coefficient of determination.

randomly initiated. In addition, the relationship between pollutant gases and $PM_{2.5}$ (PM_{10}) is not exact. Finally, other factors contributing to $PM_{2.5}$ (PM_{10}), such as water-soluble ions, are not considered. Therefore, these two methods were repeated 10 times, and each dataset was randomly divided into two parts (three-fourth samples and one-fourth samples) for training models and validating the fitting degree of models. Because the current study aims to study the relationship between pollutant gases and $PM_{2.5}$ (PM_{10}), which is not consistent with time-series prediction, we randomly divided the datasets. Moreover, we hope to construct a more diverse dataset to observe whether GEP could model the trend of increase or decrease in $PM_{2.5}$ and PM_{10} concentrations. The maximum, minimum, and mean values of fitting degrees are shown in Table 1 so that the rough effect power of pollutant gases can be explained clearly. Moreover, the results with the highest fitting result out of 10 repeated experiments are shown in Figures 4–7, where the computational values obtained with the trained model (the function model from GEP and the network model from the BP neural network) and the

TABLE 2 RMSE for GEP and BP neural networks.

Dataset	PM_{10}	
	RMSE	RMSE
Monitoring site 1	8.7365	13.2739
Monitoring site 2	9.0657	14.7867
Monitoring site 3	10.7637	45.8769
Monitoring site 4	12.7653	19.7682
Monitoring site 5	15.8766	23.7681
Average value of the whole city	14.6438	25.6792
Dataset	PM_{10}	
	RMSE	RMSE
Monitoring site 1	13.8741	29.7327
Monitoring site 2	14.5346	52.8653
Monitoring site 3	17.8643	49.7824
Monitoring site 4	23.8675	41.8654
Monitoring site 5	34.7682	32.8503
Average value of whole city	21.7357	36.5908

observational values are displayed in Figures 4–7 so that the fitting degree can be illustrated directly with these figures. The difference between Figures 4, 5 is obvious, namely, the performance of GEP is better than that of the BP neural networks. Moreover, GEP and BP neural networks performed similarly for both PM_{10} and $PM_{2.5}$, indicating that pollutant gases did not contribute much more to PM_{10} than $PM_{2.5}$ (Figure 4 *versus* Figures 5, 6 *versus* Figure 7). Similarly, BP neural networks performed no better than GEP for modeling the relationship between pollutant gases and PM_{10}. Although GEP cannot completely fit the concentration of $PM_{2.5}$ (PM_{10}) with pollutant gases, the trend of $PM_{2.5}$ (PM_{10}) can be constructed (i.e., increase or decrease). These formulas can be used to analyze the hidden laws about $PM_{2.5}$ and PM_{10} with mathematical modeling and help predict concentrations of $PM_{2.5}$ and PM_{10}.

Discussion

Experimental results show that pollutant gases influence $PM_{2.5}$ concentrations more seriously than PM_{10}. The results from GEP show that the average influence power of pollutant gases on $PM_{2.5}$ ranges from 0.2711 to 0.6359 and the average influence power of pollutant gases on PM_{10} ranges from -0.3231 to 0.2242. The results returned by the BP neural network indicate that the average influence power of pollutant gases on $PM_{2.5}$ ranges from −0.0704 to 0.1979 and the average influence power of pollutant gases on PM_{10} ranges from −0.2090 to 0.1989. There are a variety of mathematical operations and formulas, such as cosine and sine functions, that could facilitate the fitting between pollutant gases and $PM_{2.5}$ (PM_{10}). At the same time, BP neural networks are easily over-fitting, and their interpretability is not better than that of GEP. The disadvantage of GEP is that it is computation-intensive; that is, the modeling process takes significantly longer than BP neural networks. The performance of GEP and BP neural networks was also evaluated by RMSE (Doreswamy et al., 2020), which is shown in Table 2. Concretely, after GEP and BPNN construct the relationship between pollutant gases and PM2.5 (PM10), the testing data could be fed into the model (GEP formula or BPNN model) to obtain the computational values of PM2.5 (PM10). Then, the computational values could be compared with observational values to figure out the coefficient of determination and RMSE. A smaller RMSE indicates better regression performance, and bigger coefficients of determination indicate better regression performance. The metrics in Table 2 also show that pollutant gases are more related with $PM_{2.5}$ than PM_{10}, and the performance of GEP is better than that of BP neural networks. In addition, the formulas with highest fitting degrees obtained by GEP are shown in Table 3.

Then, linear regression was applied to study which pollutant gases play more important roles than others in affecting the concentration of $PM_{2.5}$ (PM_{10}). The influence power of each type of pollutant gas on $PM_{2.5}$ (PM_{10}) can be signified by the coefficients of each item (each pollutant gas), which are shown in Table 4 in descending order. Table 4 demonstrates that the pollutant gas which is more related to $PM_{2.5}$ (PM_{10}) is different for different monitoring sites. Compared with the average concentrations of O_3 in 1 hour and 8 hours, CO and contribute to $PM_{2.5}$ more, and NO_2 and SO_2 are more relevant to PM_{10} overall.

Compared with relevant research studies on the prediction of $PM_{2.5}$ (PM_{10}) concentrations, we summarized the key information of different algorithms in Table 5. The current study did not adopt pre-processing procedures for raw data, algorithms, or evaluation metrics. The algorithms in the literature include time-series regression, which applied post-temporal $PM_{2.5}$ (PM_{10}) data to predict current $PM_{2.5}$ (PM_{10}), and regression, which applied the other factors that influence $PM_{2.5}$ (PM_{10}) to model the relationship between them. Artificial intelligence (AI) can be used to predict the concentration and the main cause of $PM_{2.5}$ (PM_{10}) to control the air quality index with the assistance of social monitoring data (vehicle, social events, meteorological factors, etc.). Moreover, the early warning for $PM_{2.5}$ (PM_{10}) could be used to improve the healthcare conditions of people. Consistent with the literature (Wang P. et al., 2015; Song et al., 2015), CO and SO_2 are the pollutant gases that are related most with $PM_{2.5}$ and PM_{10}. The main sources of CO and SO_2 are automobile exhaust and fossil fuel combustion. Therefore, we should encourage clean energy for household and industrial use. Moreover, the dataset is from 1 January to 26 April, which has cold temperatures, so the main reason for pollution is heating in the winter (the heating season in Xi'an spans from 15 November to 15 March). Because we only include the pollutant gas as an independent variable in the model, no further conclusion comes into being.

Hypothesis and limitations

The present research only considers the effect of pollutant gases on $PM_{2.5}$ (PM_{10}); there are many other factors such as

TABLE 3 Formulas with best fitting degrees using GEP.

$PM_{2.5}$	
Dataset	Formula
Monitoring site 1	$5^*x_3 + x_5/x_1 + 3^* \tan(\sinh(2^*x_2)) - x_4$
Monitoring site 2	$3^*x_1 - 3^* \ln(x_3) + \max(x_2, abs(-x_4)) + \max(x_2, abs(x_4)) - 9^*x_4 + abs(-x_4)$
Monitoring site 3	$2^* \log_2(\sinh(x_3)) - (x_5 + abs(-(1/x_4^2 + \max(sqrt(x_1), x_2)))) + 3^*x_2$ $-(x_4 + abs(x_4 + abs(-x_2))) - (x_5 + sqrt(x_5 + \log_2(x_1)))$
Monitoring site 4	$-3^* \log_2(\exp(\tan(-\log_2(x_3)))) - \log_2(\exp(x_5))$ $+\cosh(\log_2(x_2)) + \cosh(3^* \log_{10}(x_1))$
Monitoring site 5	$sqrt(x_1) + 3^*x_3 - 9^* \max(x_5, \max(x_4, x_5))$ $+3^*x_1 + sqrt(x_2) + \max(3^*x_4, x_3)$
Average value of the whole city	$2^*x_2 - x_5 - \max(x_4, x_5) - x_3 + x_3^* \ln(x_1)$
$PM_{2.5}$	
Dataset	Formula
Monitoring site 1	$3^* \log_2(1/x_2) + 6^*x_3 + 1/x_2 + 18^* \log_2(1/x_3)$
Monitoring site 2	$(x_4/\log_{10}(x_4)^{2^*}(abs(x_2))^2 + (2x_1/(-x_1))/2 + (abs(\log_{10}(x_2))/x4)^2 \log_{10}(x_2))/x4)^2$ $+3^*x_2 + \log 2(x_4)^{2^*}(\log_2(x_2) - x_4) + 2^*x_1/(-x_2)/2$
Monitoring site 3	$\log_{10}(\log_{10}(x_3)^*x_1)^*x_5 + \max(x_5, (x_2^* \log_{10}(\log_{10}(x_3)^* \min(x_1, x_3))))$ $+x_2^* \log_{10}(\log_{10}(x_3)^* \min(x_1, x_5)) + 2^* \log_{10}(\max(3^* \max(x_4, x_2), x_5^2))$ $^*(x_5 + x_1) + \log_{10}(\cosh(x_1))^*(-x_5)$
Monitoring site 4	$(x_2 - (\log_2(\log_2(x_3)) + x_5)/2)^* \log_{10}(x_1) + \log_{10}((x_3 + x_2)/2^*x_1/x_2)$ $^*(x_1 - x_2) + abs(\min(\tan((x_5 + x_2)/2), x_2))^* \log 2(x_3)$ $+abs(\min(x_1, x_1^2x_2))^* \log_2(x_2) - x_2 - x_5$
Monitoring site 5	$1/(x_4/(x_5/abs(x_1))) + (x_3 - x_1) + 2^*(1/(x_4/(x_5/x_1)) + (x_3 - x_1))$ $-4^*(x_2/x_1{}^*x_5) + x_4/(x_1 - 1)^*x_3$
Average value of the whole city	$-2^*x_3 + 2^* \cosh(sqrt(2^*x_1)) - x_1 - 4^*x_4$ $+\cosh(\log_2(x_2)) + \cosh(sqrt((\ln(x_3) + \min(x_1, x_4))/2))$

(Note: x_1–x_5 refer to SO_2, NO_2, CO, average concentration of O_3 in 1 hour, and average concentration of O_3 in 8 hours, respectively).

TABLE 4 Correlation degrees of five pollutant gases with $PM_{2.5}$ and PM_{10}.

Dataset	$PM_{2.5}$	PM_{10}
Monitoring site 1	CO, SO_2, average concentration of O_3 in 8 hours, NO_2, and average concentration of O_3 in 1 hour	CO, average concentration of O_3 in 1 hour, average concentration of O_3 in 8 hours, SO_2, and NO_2
Monitoring site 2	Average concentration of O_3 in 8 hours, SO_2, CO, NO_2, and average concentration of O_3 in 1 hour	SO_2, NO_2, average concentration of O_3 in 8 hours, CO, and average concentration of O_3 in 1 hour
Monitoring site 3	CO, NO_2, average concentration of O_3 in 1 hour, SO_2, and average concentration of O_3 in 8 hours	NO_2, CO, average concentration of O_3 in 8 hour, SO_2, and average concentration of O_3 in 1 hour
Monitoring site 4	SO_2, average concentration of O_3 in 1 hour, NO_2, CO, and average concentration of O_3 in 8 hours	NO_2, average concentration of O_3 in 1 hour, SO_2, CO, and average concentration of O_3 in 8 hour
Monitoring site 5	SO_2, CO, NO_2, average concentration of O_3 in 8 hours, and average concentration of O_3 in 1 hour	NO_2, SO_2, average concentration of O_3 in 8 hours, CO, and average concentration of O_3 in 1 hour
Average value of whole city	CO, NO_2, SO_2, average concentration of O_3 in 1 hour, and average concentration of O_3 in 8 hours	NO_2, SO_2, CO, average concentration of O_3 in 8 hour, and average concentration of O_3 in 1 hours

meteorological factors, human behavior, and chemical reactions that could be considered together. In addition, GEP is computation-intensive, requiring a significant amount of time. Moreover, the modeling methods should be improved, such as deep learning (Li et al., 2023), and applied to the relevant topics. Although the interpretability is improved by GEP, the formulas are not consistent with human thoughts; some novel methods could also be applied to show a more direct relationship, such as

TABLE 5 Key points in relevant research studies.

Reference	Pre-processing	Data (place)	Algorithm	Evaluation metrics	Result
Prediction of daily mean and 1-h maximum PM2.5 concentrations and applications in Central Mexico using satellite-based machine-learning models (Gutiérrez-Avila et al., 2022)	N.A.	Geographical location, date, meteorological factors, and satellite data, The height of the planetary boundary layer (Mexico)	Extreme gradient boosting (XGBoost)	Mean absolute errors (MAE)	3.68
PM2.5 analog forecast and Kalman filter post-processing for the community Multiscale Air Quality (CMAQ) model (Djalalova et al., 2015)	Outlier value is rejected, such as PM2.5 > 500 and the incorrect value	The observational PM2.5 dataset consists of 716 monitoring sites found in the AirNow data set. (United States of America)	Kalman filter post-processing	MAE and correlation coefficient	50%–75% and 40%–60%
Forecasting air pollution particulate matter (PM2.5) Using machine learning regression models (Doreswamy et al., 2020)	Fill missing data	Taiwan Air Quality Monitoring Network (TAQMN) dataset available for 76 stations in different locations: geographical data, chronological data. and meteorological data (Taiwan)	Linear regression and random forest Regressor, gradient boosting regressor, k-neighbors regressor, MLP regressor, and decision tree regressor (CART)	RMSE, MAE, mean square error (MSE), and coefficient of determination (R^2)	0.8891, 0.0169 0.1302 0.0380 (best model)
Machine learning-based model to estimate PM2.5 concentration levels in Delhi's atmosphere (Kumar et al., 2020)	N.A.	Various atmospheric and surface factors such as wind speed, atmospheric temperature pressure, etc. (Indian)	Extra-trees regressor algorithm with AdaBoost	MAE, RMSE, and R^2	14.79, 25.11, and 92.96%
An improved deep learning model for predicting daily PM2.5 concentration (Xiao et al., 2020)	Fill missing data and data normalization, date one-hot encoding	Satellite data and meteorological data (China)	Weighted long short-term memory Neural network extended model (WLSTME)	RMSE and MAE	40.67 and 26.10
A hybrid land use regression/ AERMOD model for predicting intra-urban variation in PM2.5 (Michanowicz et al., 2016)	AERMOD preprocessing	Meteorological data (United States of America)	Land use regression	RMSE	1.34 (summer) and 1.43 (winter)
Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China (Pak et al., 2020)	Identifying the inherent interaction between the given variables with mutual information-based spatiotemporal correlation analysis	Air quality and meteorological data (China)	CNN–LSTM (convolutional neural networks–long short-term memory networks)	RMSE and MAE	5.357 and 4.971
Machine learning and deep learning modeling and simulation for predicting PM2.5 concentrations (Peng et al., 2022)	N.A.	Meteorological data (China)	Random forest and XGBoost	R^2	0.761
Combining machine learning and numerical simulation for high-resolution PM2.5 concentration forecast (Bi et al., 2022)	N.A.	Meteorological data and land use data (China)	Random forest	RMSE, mean absolute percentage error (MAPE), and R^2	16.7, 34.3, and 0.76

fuzzy cognitive maps (Zhang et al., 2019). Moreover, the computational resource and data resource requirement (large datasets covering more factors and more regions) impose significant burden for practice.

Conclusion

GEP was employed to model the impact of pollutant gases on concentrations of $PM_{2.5}$ (PM_{10}); the influence power is measured with the coefficient of determination. BP neural networks were used as the baseline method. Experimental results show that the influence power of pollutant gases on $PM_{2.5}$ and PM_{10} is between -0.0704 and 0.6359 and between -0.3231 and 0.2242), respectively. The performance of the models is also compared with RMSE (root mean squared error) (Doreswamy et al., 2020). GEP achieved an RMSE of [8.7365–14.6438] for $PM_{2.5}$ and the RMSE of [13.2739–45.8769] for PM_{10}, and BP neural networks achieved the average RMSE of [13.8741–34.7682] for $PM_{2.5}$ and the RMSE of [29.7327–52.8653] for PM_{10}. For the coefficient of determination, GEP and BPNN achieved mean 0.2091–0.6539 and -0.0704–0.1979 ($PM_{2.5}$) and mean -0.3231–0.2242 and -0.1105–0.1989 (PM_{10}). GEP

achieved better RMSE and coefficient of determination metrics than the BPNN. The results from GEP are more explainable than those from the BPNN because the formula could directly reflect the correlation between independent variables (pollutant gas) and dependent variables ($PM_{2.5}$/PM_{10}). The formulas obtained with GEP can be applied to study carefully to draw more conclusions from every angle. The heterogeneous relationship modeled by GEP in different seasons or specific regions could be used to monitor the causality of $PM_{2.5}$ and PM_{10} so that pollution could be restricted. Then, results show that $PM_{2.5}$ is more correlated to CO, whereas PM_{10} is more correlated to NO_2 and SO_2, which is inferred using linear regression. Above methods and relevant conclusions can be beneficial in controlling and forecasting $PM_{2.5}$ (PM_{10}) concentrations. Although some conclusions came into being, there are still some problems to be solved in the future, such as some negative values that will certainly not exist, which can be tackled by correcting the unreasonable chromosomes in GEP, improving the mechanism of the neural network, adjusting proper algorithm parameters for different datasets, or adding more attributes that affect $PM_{2.5}$ (PM_{10}) to the dataset. All of the above results show that GEP can be applied in environmental modeling to get more quantitative and explainable conclusions.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found here: The dataset is available: https://www.aqistudy.cn/historydata/.

Author contributions

XW: investigation, methodology, software, supervision, validation, visualization, and writing–original draft. KZ: conceptualization, formal analysis, methodology, project administration, resources, and writing–original draft. PH: project administration, resources, visualization, and writing–review and editing. MW: formal analysis, funding acquisition, visualization, and writing–review and editing. XL: data curation, investigation, methodology, and writing–review and editing. YZ: project administration, supervision, validation, and writing–review and editing. QP: investigation, methodology, validation, visualization, and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The Project Supported by the Natural Science Basic Research Plan in Shaanxi Province of China (2022JQ-175) and the Scientific Research Plan of Shaanxi Education Department (22JK0303).

Conflict of interest

Author KZ was employed by Chongqing Chang'an Industrial Co., Ltd. Author PH was employed by Shenzhen Metro Operation Group Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Apte, J. S., Marshall, J. D., Cohen, A. J., and Brauer, M. (2015). Addressing global mortality from ambient PM2.5. *Environ. Sci. Technol.* 49, 8057–8066. doi:10.1021/acs.est.5b01236

Azamathulla, H. M. (2012). Gene-expression programming to predict scour at a bridge abutment. *J. Hydroinformatics* 14, 324–331. doi:10.2166/hydro.2011.135

Bi, J., Knowland, K. E., Keller, C. A., and Liu, Y. (2022). Combining machine learning and numerical simulation for high-resolution PM2.5 concentration forecast. *Environ. Sci. Technol.* 56, 1544–1556. doi:10.1021/acs.est.1c05578

Bossmann, K., Bach, S., Höflich, C., Valtanen, K., Heinze, R., Neumann, A., et al. (2016). Holi colours contain PM10 and can induce pro-inflammatory responses. *J. Occup. Med. Toxicol.* 11, 42. doi:10.1186/s12995-016-0130-9

Djalalova, I., Delle Monache, L., and Wilczak, J. (2015). PM2.5 analog forecast and Kalman filter post-processing for the community multiscale air quality (CMAQ) model. *Atmos. Environ.* 108, 76–87. doi:10.1016/j.atmosenv.2015.02.021

Doreswamy, K. S. H., Km, Y., and Gad, I. (2020). Forecasting air pollution particulate matter (PM2.5) using machine learning regression models. *Procedia Comput. Sci.* 171, 2057–2066. doi:10.1016/j.procs.2020.04.221

Ferreira, C. (2001a). Gene expression programming: a new adaptive algorithm for solving problems. *Comput. Sci.*, 87–129. doi:10.48550/arXiv.cs/0102027

Ferreira, C. (2001b). *Gene expression programming: a new adaptive algorithm for solving problems*. arXiv preprint cs/0102027.

Francesca Dominici, M. G., Sunstein, C. R., and Sunstein, C. R. (2014). Particulate matter matters. *Science* 344, 257–259. doi:10.1126/science.1247348

Frank, A., Fabregat-Traver, D., and Bientinesi, P. (2015). Large-scale linear regression: development of high-performance routines. *Appl. Math. Comput.* 275, 411–421. doi:10.1016/j.amc.2015.11.078

Gautam, S., Yadav, A., Tsai, C. J., and Kumar, P. (2016). A review on recent progress in observations, sources, classification and regulations of PM 2.5 in Asian environments. *Environ. Sci. Pollut. Res. Int.* 23, 21165–21175. In Press. doi:10.1007/s11356-016-7515-2

Gutiérrez-Avila, I., Arfer, K. B., Carrión, D., Rush, J., Kloog, I., Naeger, A. R., et al. (2022). Prediction of daily mean and one-hour maximum PM2.5 concentrations and applications in Central Mexico using satellite-based machine-learning models. *J. Expo. Sci. Environ. Epidemiol.* 32, 917–925. doi:10.1038/s41370-022-00471-4

Jerrett, M., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., et al. (2009). Long-term ozone exposure and mortality. *N. Engl. J. Med.* 360, 1085–1095. doi:10.1056/nejmoa0803894

Kicsiny, R. (2014). Multiple linear regression based model for solar collectors. *Sol. Energy* 110, 496–506. doi:10.1016/j.solener.2014.10.003

Kumar, S., Mishra, S., and Singh, S. K. (2020). A machine learning-based model to estimate PM2.5 concentration levels in Delhi's atmosphere. *Heliyon* 6, e05618. doi:10.1016/j.heliyon.2020.e05618

Li, J., Crooks, J., Murdock, J., de Souza, P., Hohsfield, K., Obermann, B., et al. (2023). A nested machine learning approach to short-term PM2.5 prediction in metropolitan areas using PM2.5 data from different sensor networks. *Sci. Total Environ.* 873, 162336. doi:10.1016/j.scitotenv.2023.162336

Liu, S., Hou, Z., and Yin, C. (2015). Data-driven modeling for UGI gasification processes via an enhanced genetic BP neural network with link switches. *IEEE Trans. Neural Netw. Learn. Syst.*

Ma, Y., Chen, R., Pan, G., Xu, X., Song, W., Chen, B., et al. (2011). Fine particulate air pollution and daily mortality in Shenyang, China. *Sci. Total Environ.* 409, 2473–2477. doi:10.1016/j.scitotenv.2011.03.017

Michanowicz, D. R., Shmool, J. L. C., Tunno, B. J., Tripathy, S., Gillooly, S., Kinnee, E., et al. (2016). A hybrid land use regression/AERMOD model for predicting intra-urban variation in PM2.5. *Atmos. Environ.* 131, 307–315. doi:10.1016/j.atmosenv.2016.01.045

Mostafa, M. M., and El-Masry, A. A. (2016). Oil price forecasting using gene expression programming and artificial neural networks. *Econ. Model.* 54, 40–53. doi:10.1016/j.econmod.2015.12.014

Nel, A. (2005). Air pollution-related illness: effects of particles. *Science* 308, 804–806. doi:10.1126/science.1108752

Ostro, B., Broadwin, R., Green, S., Feng, W. Y., and Lipsett, M. (2006). Fine particulate air pollution and mortality in nine California counties: results from CALFINE. *Env. Health Perspect.* 114 (1), 29–33. doi:10.1289/ehp.8335

Özcan, F. (2012). Gene expression programming based formulations for splitting tensile strength of concrete. *Constr. Build. Mater.* 26, 404–410. doi:10.1016/j.conbuildmat.2011.06.039

Pak, U., Ma, J., Ryu, U., Ryom, K., Juhyok, U., Pak, K., et al. (2020). Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: a case study of Beijing, China. *Sci. Total Environ.* 699, 133561. doi:10.1016/j.scitotenv.2019.07.367

Peng, J., Han, H., Yi, Y., Huang, H., and Xie, L. (2022). Machine learning and deep learning modeling and simulation for predicting PM2.5 concentrations. *Chemosphere* 308, 136353. doi:10.1016/j.chemosphere.2022.136353

Pui, D. Y. H., Chen, S. C., and Zuo, Z. (2014). PM 2.5 in China: measurements, sources, visibility and health effects, and mitigation. *Particuology* 13, 1–26. doi:10.1016/j.partic.2013.11.001

Schmidt, M., and Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science* 324, 81–85. doi:10.1126/science.1165893

Song, Y.-Z., Yang, H.-L., Peng, J.-H., Song, Y.-R., Sun, Q., and Li, Y. (2015). Estimating PM2. 5 concentrations in Xi'an city using a generalized additive model with multi-source monitoring data. *PloS one* 10, e0142149. doi:10.1371/journal.pone.0142149

Sun, Z., An, X., Tao, Y., and Hou, Q. (2013). Assessment of population exposure to PM10 for respiratory disease in Lanzhou (China) and its health-related economic costs based on GIS. *BMC Public Health* 13, 891. doi:10.1186/1471-2458-13-891

Tosun, E., Aydin, K., and Bilgili, M. (2016). Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures. *Alexandria Engineering Journal* 55, 3081–3089. doi:10.1016/j.aej.2016.08.011

Wang, G., Su, Y., and Shu, L. (2016). One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. *Renew. Energy* 96, 469–478. doi:10.1016/j.renene.2016.04.089

Wang, L., Ren, T., Nie, B., Chen, Y., Lv, C., Tang, H., et al. (2015a). Development of a spontaneous combustion TARPs system based on BP neural network. *Int. J. Min. Sci. Technol.* 25, 803–810. doi:10.1016/j.ijmst.2015.07.016

Wang, P., Cao, J., Tie, X., Wang, G., Li, G., Hu, T., et al. (2015b). Impact of meteorological parameters and gaseous pollutants on PM2. 5 and PM10 mass concentrations during 2010 in Xi'an, China. *Aerosol Air Qual. Res.* 15, 1844–1854. doi:10.4209/aaqr.2015.05.0380

Wang, Y., Lu, C., and Zuo, C. (2015c). Coal mine safety production forewarning based on improved BP neural network. *Int. J. Min. Sci. Technol.* 25, 319–324. doi:10.1016/j.ijmst.2015.02.023

Wu, C. H., Lin, I. S., Wei, M. L., and Cheng, T. Y. (2013). Target position estimation by genetic expression programming for mobile robots with vision sensors. *IEEE Trans. Instrum. Meas.* 62, 3218–3230. doi:10.1109/tim.2013.2272173

Xiao, F., Yang, M., Fan, H., Fan, G., and Al-qaness, M. A. A. (2020). An improved deep learning model for predicting daily PM2.5 concentration. *Sci. Rep.* 10, 20988. doi:10.1038/s41598-020-77757-w

Yang, Z., Chen, Y., Tang, Z., and Wang, J. (2016). Surface EMG based handgrip force predictions using gene expression programming. *Neurocomputing* 207, 568–579. doi:10.1016/j.neucom.2016.05.038

Yassin, M. A., Alazba, A. A., and Mattar, M. A. (2016). A new predictive model for furrow irrigation infiltration using gene expression programming. *Comput. Electron. Agric.* 122, 168–175. doi:10.1016/j.compag.2016.01.035

Yu, F., and Xu, X. (2014). A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. *Appl. Energy* 134, 102–113. doi:10.1016/j.apenergy.2014.07.104

Zhang, K., Pan, Q., Yu, D., Wang, L., Liu, Z., Li, X., et al. (2019). Systemically modeling the relationship between climate change and wheat aphid abundance. *Sci. Total Environ.* 674, 392–400. doi:10.1016/j.scitotenv.2019.04.143

Zhou, J., Wan, X., Zhang, J., Yan, Z., and Li, Y. (2015). Modeling of constitutive relationship of aluminum alloy based on BP neural network model. *Mater. Today Proc.* 2, 5023–5028. doi:10.1016/j.matpr.2015.10.092

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 03 September 2024
DOI 10.3389/fenvs.2024.1440396

OPEN ACCESS

EDITED BY
Jawad Ahmad,
Swedish College of Engineering and Technology, Pakistan

REVIEWED BY
R Suresh,
M S Ramaiah University of Applied Sciences, India
Harish Kumar Muniswamy,
Christ University, India

*CORRESPONDENCE
Satish Kumar,
satishkumar.vc@gmail.com
Ketan Kotecha,
director@sitpune.edu.in

RECEIVED 29 May 2024
ACCEPTED 06 September 2024
PUBLISHED 03 September 2024

CITATION
Joshi DD, Kumar S, Patil S, Kamat P, Kolhar S and Kotecha K (2024) Deep learning with ensemble approach for early pile fire detection using aerial images.
Front. Environ. Sci. 12:1440396.
doi: 10.3389/fenvs.2024.1440396

Deep learning with ensemble approach for early pile fire detection using aerial images

Dhyey Divyeshkumar Joshi[1,2], Satish Kumar[1,3]*, Shruti Patil[1,3], Pooja Kamat[3], Shrikrishna Kolhar[3] and Ketan Kotecha[1,3]*

[1]Symbiosis Centre of Applied Artificial Intelligence, Symbiosis International (Deemed University), Pune, India, [2]Devang Patel Institute of Advance Technology and Research, Charusat University, Anand, Gujarat, India, [3]Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune, Maharashtra, India

Wildfires rank among the world's most devastating and expensive natural disasters, destroying vast forest resources and endangering lives. Traditional firefighting methods, reliant on ground crew inspections, have notable limitations and pose significant risks to firefighters. Consequently, drone-based aerial imaging technologies have emerged as a highly sought-after solution for combating wildfires. Recently, there has been growing research interest in autonomous wildfire detection using drone-captured images and deep-learning algorithms. This paper introduces a novel deep-learning-based method, distinct in its integration of infrared thermal, white, and night vision imaging to enhance early pile fire detection, thereby addressing the limitations of existing methods. The study evaluates the performance of machine learning algorithms such as random forest (RF) and support vector machines (SVM), alongside pre-trained deep learning models including AlexNet, Inception ResNetV2, InceptionV3, VGG16, and ResNet50V2 on thermal-hot, green-hot, and white-green-hot color images. The proposed approach, particularly the ensemble of ResNet50V2 and InceptionV3 models, achieved over 97% accuracy and over 99% precision in early pile fire detection on the FLAME dataset. Among the tested models, ResNet50V2 excelled with the thermal-fusion palette, InceptionV3 with the white-hot and green-hot fusion palettes, and VGG16 with a voting classifier on the normal spectrum palette dataset. Future work aims to enhance the detection and localization of pile fires to aid firefighters in rescue operations.

KEYWORDS

aerial imaging1, deep learning2, fire piles monitoring3, thermal infrared images4, RGB images5, wildfire detection system6

1 Introduction

Wildfires have recently caused significant damage to forests, wildlife habitats, farms, residential areas, and entire ecosystems. Data from the National Interagency Fire Center (NIFC) reveals that between 2012 and 2021 in the United States, 22,897 fires ravaged more than 1,082,125 acres, resulting in over $7.4 billion in losses. Fires are rare but can quickly lead to extensive harm and property damage. The National Fire Protection Association (NFPA) predicts that in 2022, the United States Emergency Fire Service will respond to approximately 1,089,844 fires, leading to 1,149 civilian deaths and an estimated $18 billion in property damage (National Fire News), (Statistics). Early identification of fires without

false alarms was critical to reduce such catastrophes. As a result, various active fire protection methods were created and are widely used in the current world. There were several methods available for detecting and monitoring fires. Satellite imagery, remotely piloted vehicles (RPV), and sensors were examples of this. These systems, however, do not enable real-time fire detection and monitoring. Some disadvantages of these technologies include: i) fires were not noticed in their early stages, ii) satellites have larger time delays, and iii) sensors become an infeasible option due to their limited range. Even if spectral imaging is employed, it may not provide an accurate result at night or in bad weather. To fight against fire at the right moment, individuals were now engaged to watch the forest, risking their lives. These troubling facts inspire academics to develop innovative early pile burn identification and monitoring methods.

In particular, technical innovations in airborne surveillance technologies can provide early rescuers and tactical soldiers with more precise fire behavioral information for improved fire catastrophe prevention and mitigation. Using computer vision and remote tracking imagery, remotely piloted aircraft proved highly effective for real-time fire monitoring and detection. Unmanned aerial vehicles (UAVs), also known as drones, offer a quicker, more agile, and cost-effective method for forest surveillance. Enhancing UAVs with remote sensors can significantly improve current techniques. Additionally, UAVs can operate in hazardous areas that are unsafe for human access. Traditional methods of finding and monitoring flames included stationing personnel in observation stations or sending choppers or corrected airborne vehicles to surveil fires using long-wave thermal vision (Habiboğlu et al., 2012, Rafiee et al., 2011, Jadon et al., 2019). Current research has examined wireless mesh network-based Internet of Things (IoT) advancements, but such infrastructures would require additional funding and verification before providing meaningful information. Satellite data was frequently used for international fire assessment at larger scales, albeit at relatively low resolution and with repeat imagery availability controlled by satellite circular trends (Afghah et al., 2018a).

Given these systems' limits and problems, the use of Unmanned Drones for fire monitoring has grown in recent decades. UAVs provide innovative characteristics and benefits such as quick deployment, mobility, a broader and adjustable field of view, and little operator engagement (Afghah et al., 2018b, Aggarwal et al., 2020, Afghah et al., 2019). UAVs can offer the following advantages: they cover a large area, even in overcast conditions, work at all hours of the day and night, are readily recovered, and are reasonably affordable, electric UAVs are also beneficial to the environment, transport various payloads for various missions and effectively cover target area (Casbeer et al., 2006, Yuan et al., 2015a) and, most critically, missions may be completed automatically without or with minimal human pilot/operator interaction. Drone technology development has increased in recent years. Nowadays, unmanned aerial vehicles (UAVs) are employed in emergency relief scenarios and activities such as forest fires and flooding, especially as a short-term solution whenever domestic connections crash due to destroyed infrastructure, miscommunication, or bandwidth limits (Shamsoshoara et al., 2020a, Shamsoshoara, 2019).

Artificial intelligence (AI) and machine learning advancements in recent years have boosted the success of image categorization, detection, legitimate forecasting, and other applications. Artificial intelligence (AI) can be utilized in numerous fields. The advancement of nanotechnology in semiconductor materials has led to the creation of advanced tensor and graphical processing units, which offer exceptional processing power for data-driven approaches. Moreover, modern drones and aerial vehicles can be equipped with compact edge technology TPU and GPU units for real-time processing, enabling early detection of fires and preventing major disasters (NVIDIA Jetson Nano Developer Kit), (Edge TPU).

Several guided learning algorithms rely on massive training sets to produce a suitably accurate model (Wu et al., 2020). Using a public fire dataset, their research detected fires using pre-trained convolutional neural networks such as mobile and AlexNet architecture. That dataset, however, was based on images of the fire acquired from the ground. To our knowledge, the fire luminosity airborne-based machine learning evaluation (FLAME) dataset was the only airborne imagery collection for pile-fire research that could be used to build fire modeling or advanced analytics systems for airborne surveillance systems. The authors used their own gathered normal-spectrum fire dataset to detect fire using XceptionNet's pre-trained convolutional neural network (CNN) architectures (Shamsoshoara et al., 2021).

In this article, researchers have utilized the FLAME dataset for pile burns. It is a drone-captured pile of fire clips and photos taken during organized slashed piles in North Arizona. Pile burns were extremely useful for studying spot fires and early-stage fires. Following thinning and restoration initiatives, forest management generally uses pile fires to clear up forest leftovers, such as branches and leaves. Various UAVs took the imagery with distinct viewpoints, magnification levels, and camera kinds, including conventional and infrared cameras. Forestry interventions were an essential strategy for reducing biofuels, and blazing chopped piles were often the most cost-effective and secure technique for re-moving gash. Large burning regions were often hard to manage, resulting in massive deforestation from wildfires. As a result, any controlled burnings were accompanied by comprehensive forest management strategies to avoid fire. Slash or debris burning was common for certain farms and families to burn leaves, crop remnants, and other agricultural trash. This method was both faster and handier than taking them away. However, being near woods increases the danger of wildfires and affects the environment, not to mention that it was outlawed in many nations and states across the United States. Using automated aerial surveillance devices can significantly decrease the effort of forest managers. Google's Federated Transfer Learning (FTL) is highlighted as a promising decentralized learning methodology that can be integrated with UAV-based fire detection systems, enhancing the efficiency and accuracy of models operating with limited resources. (Bonawitz et al., 2019). Portable technologies, such as unmanned aerial vehicles (drones), employ localized information to train a model with raw data the drone collects. They then communicate anticipated data with the FTL. Later, the FTL gets all model parameters via various drones and starts combining them before releasing an optimal model to all drones so they can concur rapidly. FTL techniques provide data secrecy and confidentiality; these techniques are well suited to narrow bandwidth and portable devices with limited charging capacity. Furthermore, it exchanges model parameters

with the FTL rather than all images and videos, reducing latency output.

The current FLAME dataset was gathered utilizing several drones (Shamsoshoara et al., 2021, Shamsoshoara et al., 2020b). As a result, the FTL solutions in this data collection are vital for industrial difficulties and open issues (Brik et al., 2020). The FLAME collection contains aerial imagery and clips of two sorts of palettes: normal range spectrum palettes and thermal fusion palettes, taken by advanced high-resolution cameras attached with drones which give four output visuals such as normal-spectrum for RGB col-or-space, Fusion heat map, white-fusion, and lastly, green-fusion having thermal palettes. In this work, the authors focused on individual binary classification tasks to assess the dataset's utility for extreme fire control problems. This paper presents a comprehensive comparative study of the performance of machine learning and deep learning models on the FLAME dataset. Gabor filter and Sobel filters extract features, and RF and SVM classifiers are used to detect early pile fire detection. Transfer learning is used to design and implement deep learning models. An ensemble voting classifier is used to further improve the deep learning models' performance. The proposed deep learning technique demonstrated robust performance on the FLAME dataset, which consists of aerial imagery for pile burn surveillance, including both normal spectrum and thermal heat map frames.

2 Related work

Recent deep learning applications include detecting and classifying visual objects, audio recognition, and natural language processing. To increase performance, researchers have done several experiments on fire detection based on deep learning. Deep learning varies significantly from conventional computer vision-based fire detection. As a result, the effort required to identify the appropriate handcrafted qualities has been redirected toward establishing a strong network and preparing training data. Another distinction was that the detector/classifier may be obtained by simultaneously training with the features in the same neural network.

Frizzi et al. (2016) proposed a fire detection network based on CNN, where features were simultaneously learned by training a multilayer perceptron (MLP) neural network classifier. Zhang et al. (2016) also presented a CNN-based fire detection system that works in a chain. If a fire is discovered, a fine-grained patch classifier is employed to pinpoint the fire patches precisely. Muhammad et al., 2019 suggested a fire detection system based on a finely tuned CNN fire detector. Squeeze Net-inspired (Ye et al., 2013) optimum neural network architecture was proposed for fire identification, positioning, and cognitive comprehension of the fire situation. Saydirasulovich et al. (2023) used the YOLOv8 model to detect forest fire smoke using UAV images. Abdusalomov et al. (2023) used the Detectron2 platform based on deep-learning methods for fast and accurate forest fire detection.

A unit in the deep layer of CNN has a huge receptive field; therefore, its activation may be viewed as a feature containing a big region of context information. This was another advantage of using CNN to train features for fire detection. Shamsoshoara et al. (2021) produced the FLAME dataset for pile burns in the forest of Northern Arizona, one of the key reference papers researchers have investigated. The FLAME collection contains aerial imagery and clips of two sorts, normal range spectrum and thermal fusion spectrum, taken by advanced high-resolution cameras that provide four output visuals: normal-spectrum view, Fusion heat map view, white-hot view, and lastly, green-hot view. The video frames were employed for two purposes: fire presence or absence classification and fire detection. The authors used an artificial neural network for fire absence or presence classification and U-Net architecture for fire detection. Xuan Truong and Kim (2012) used an adaptive Gaussian mixture model to detect moving regions, a fuzzy c-means algorithm to segment fire regions, and SVM applied on parameters extracted to detect fire and non-fire regions. Rinsurongkawong et al. (2012) proposed a method for early fire detection based on the Lucas-Kanade optical flow algorithm.

3 Proposed methodology

Image identification is among the most challenging tasks in computer vision. Image identification in the past algorithms used RGB or Red and Yellow or only green channel comparison to distinguish distinct items in frames or films, such as fire.

RGB luminosity comparing methods, for example, which normally employ a specified threshold to locate the fire, may identify sunset and sunrise as false positives. A guided machine learning technique and certain advanced CNN principles are used in this work to detect pile fires in a region using drone-camera-captured footage, as shown in Figure 1. When pile-fire and no-pile-fire components coexist in a mixed image, the frame is designated "pile-fire," when there is no pile burn, the frame is labeled "no-pile-fire." The Forward-Looking-Infrared (FLIR)-Vue-Pro-R-Camera, Zen-muse X4S, and DJI Phantom 3 cameras are used to identify four different types of recorded clips, such as white-hot, green-hot, thermal heat mapping, and normal spectrum of imagery. The binary classification models employed in this study are the Resnet50V2 network, the AlexNet network, the InceptionV3 network, and a VGG-16 network integrated with an ensemble voting classifier. The selection of ResNet50V2, AlexNet, InceptionV3, and VGG-16 models was based on their proven efficacy in handling complex image classification tasks and their ability to generalize well across varied datasets, making them suitable for the challenging task of early fire detection in aerial imagery. Since pile-fire detection is a binary classification problem with two possible outcomes—pile-fire and no-pile-fire—the output layer consisted of two neurons and employed a sigmoid activation function. The sigmoid activation function is represented by the following Equation 1.

$$p\left(target\_label = pile - fire\right) = \sigma\left(target\_label = pile - fire \mid \zeta\left(\theta\right)\right) = 1/\left(1 + \hat{e}\left(-\zeta\left(\theta\right)\right)\right) \tag{1}$$

where $\zeta(\theta)$ was the output layer value obtained from the input neuron, each neuron is nothing more than a pixel value of each pixel, then summed it up with the hidden network weights. The output value of the Sigmoid activation function was the probability of pile-fire detection based on the loaded frames into the network. While training the proposed pre-trained

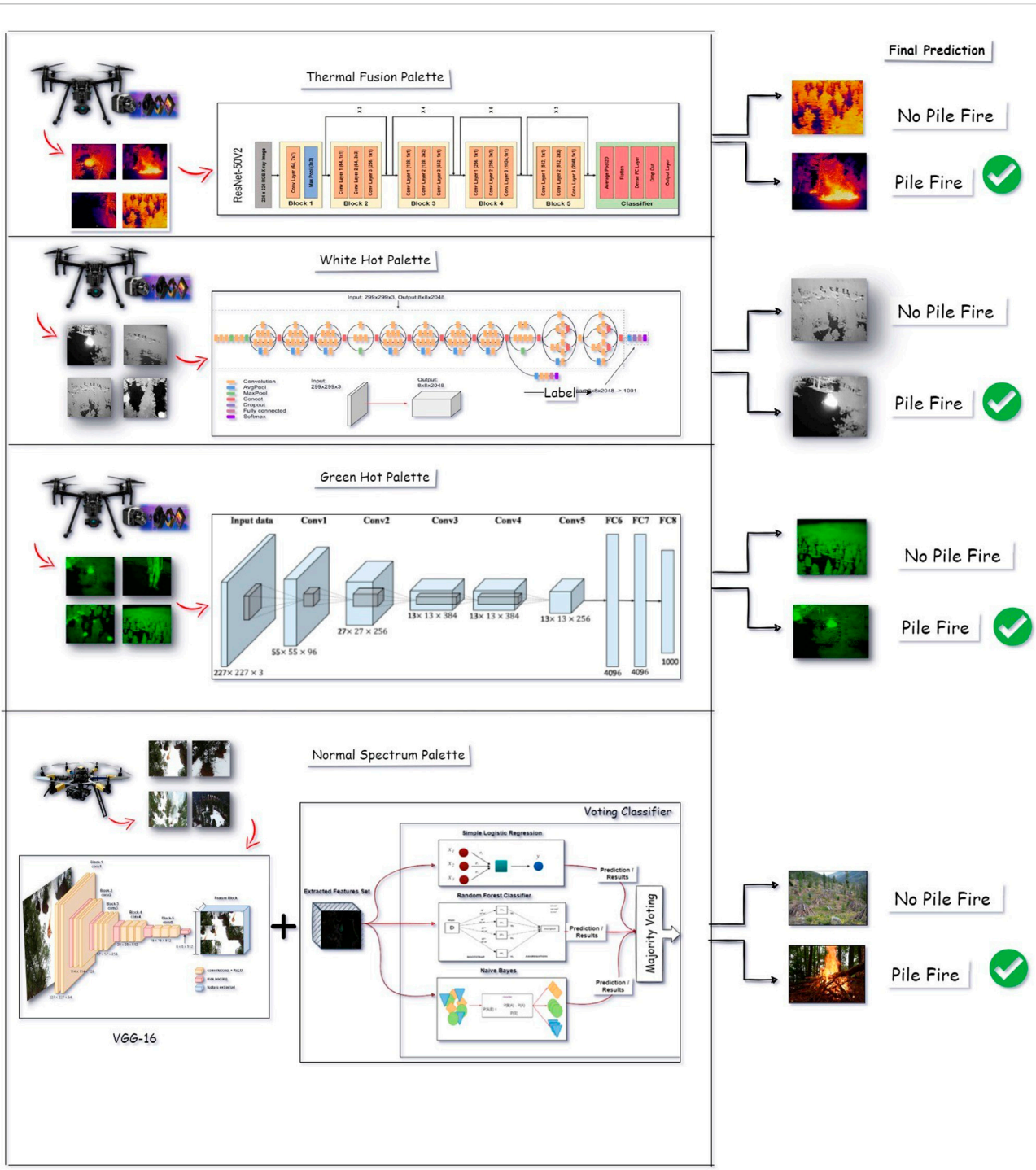


FIGURE 1
General architecture of multi-vision cameras input and different use-case for early pile fire detection.

networks, the network weights are calculated using weight initialization techniques; also, a loss function is employed using the tensor-flow library to improve network accuracy and determine the optimal weight values. Although the challenge in this part involves binary-classification, the error function studied is a binary-cross-entropy binary cross entropy which is calculated as shown in Equation 2 below:

$$LBCE\ (z, \hat{z}) = -1 \Big/ N \Sigma zi * log\ (p(\hat{zi})) + (1 - zi) * log\ (1 - p(\hat{zi})) \quad (2)$$

where N, at each epoch is the sample population in each batch utilized to adjust the gradient descent, and z is the actual true label for the frames training dataset. Classified as pile-fire (z = 1)

and z = 0, if classified as no-pile-fire. The projected probability of a frame belonging to the fire class was given by ($\hat{y}$). During the learning phase, the Adamax optimizer optimizer regulates the error function and the optimal weights on each hidden layer.

3.1 Residual neural network–Resnet-50: Version 2.0

ResNet50V2 is a typical pre-trained convolutional neural network capable of image recognition, auto-encoding, and classification. ResNet50V2 is a variant of the ResNet model with forty-eight convolutions, one max-pooling layer, and one average-pooling layer (Xuan Truong and Kim, 2012), (Rinsurongkawong et al., 2012). ResNet50V2 is an upgraded version of the ResNet50 architecture that outperformed ResNet50 and ResNet101 on the ImageNet repository. ResNet50V2 modified the transmission formulation of the linkages between blocks. ResNet50V2 creates a 10 × 10 2048 feature map from the input picture on its final feature ex-tractor layer. The residual network is generated from a 50-layer residual network architecture and includes 177 layers in total. The weights of the pre-trained network's previous layers (1–174) are frozen. The weights of several early layers can be frozen to accelerate network training and considerably prevent over-fitting to new data.

3.2 AlexNet: Version 2010

The AlexNet model has enhanced network depth compared to LeNet-5. AlexNet encompasses eight tiers of trainable parameters, as discussed in (Hu et al., 2018), (Muhammad et al., 2019). Its architectural design consists of five layers, comprising a max-pooling layer, three fully connected layers, and a final layer with an output layer employing the ReLU activation function to produce outputs. Dropout layers are also used to avoid overfitting in their model. This model takes images with dimensions 227 × 227 × 3 as input. The resultant feature vector has the dimensions (13 × 13 × 256). The number of filters rises as input images move further into the levels.

3.3 Inception-Net: Version 3.0

The InceptionV3 is a pre-trained convolution neural network used for image classification. The InceptionV3 is a modified version of the base models like the InceptionV1, launched by Google in 2014. The InceptionV3 model is a slightly advanced and optimized version of the inception-v1 model. Many techniques were used to optimize and stabilize the model (Shafiq and Gu, 2022), (Zhou et al., 2016). As a result, they created deeper networks than the Inception-v1 and v2 models, though it needed to be more effective, it took less computational power. The Inception-v3 model has forty-two layers, which is slightly higher than the previous base models. Several improvements have been made to Inception v3, such as factorization being fast as smaller convolutions, spatial factorization being converted to asymmetric convolutions, and the efficient grid size being reduced. These were the primary changes made to the InceptionV3 model.

3.4 Visual geometry group networks–VGG-Net: Version 16

The VGG-16 is also a modified version of the VGG model with sixteen convolution layers. As per the observation of the VGG-16 architecture, it contains sixteen convolutional layers. This model also has the same 3 × 3 size filters on each layer (Kundu et al., 2018), (Szegedy et al., 2017), and (Wang et al., 2017). This study used the VGG-16 model as a pre-trained model, which was trained on an ImageNet dataset. The parameters can be used as hyper-parameters to improve the accuracy. This network setup accepts a fixed size of 227 × 227 image pixels with three channels - Red, Green, and Blue. The only image preprocessing performed on the image is normalizing each pixel's RGB values. Then, the image is passed through the ReLU activation function with a stack of two convolution layers having a filter size of 3 × 3. These 2 convolution layers include 64 filters, each with stride and padding by 1. This setup is also helpful in preserving spatial resolution. In the last layer, the dimension of the output layer is kept the same as the dimensions of the input image. A window size of 2 × 2 and a stride of two pixels were applied to the images.

3.5 Ensemble vote classifier–Voting classifier

A voting classifier is an ensemble learning technique. Voting classifiers can aggregate multiple base model outputs and give the best result for a specific problem. The voting aggregation method considers each base model probability and combines them to produce an average probability of a model (Mueller et al., 2013), (Borges and Izquierdo, 2010), (Qureshi et al., 2016). There are two methods of voting criteria. Hard voting gives an integer number as a probabilistic output class, while soft voting means a floating-point-based probabilistic output class. Thus, a voting classifier can be used as a hard or soft vote for optimal results. The proposed model is trained using three basic estimators for pile fire classification: Logistic Regression, RF, and Gaussian naive Bayes as shown in Figure 2. Developers have the flexibility to select between hard and soft vote aggregators using the parameters "voting = 'soft'" or "voting = 'hard'". In the proposed model, a soft voting classifier is utilized with a weight distribution of 1:2:1, where the RF model is assigned twice the weight compared to others.

4 Dataset

This section contains information about the captured images and recorded clips captured by drones. These video clips were played at different frames-per-second speeds. To create an image (frame) dataset from a video clip, it is edited and cut into frames, where each frame may contain a black-out, pile-fired, or no-plie-fired type of background. The FLAME collection, hosted on the IEEE POST website, is freely accessible. Users can download the full high-resolution video clips and corresponding images, which are organized into four categories based on the type of drone camera and palette used. The dataset is

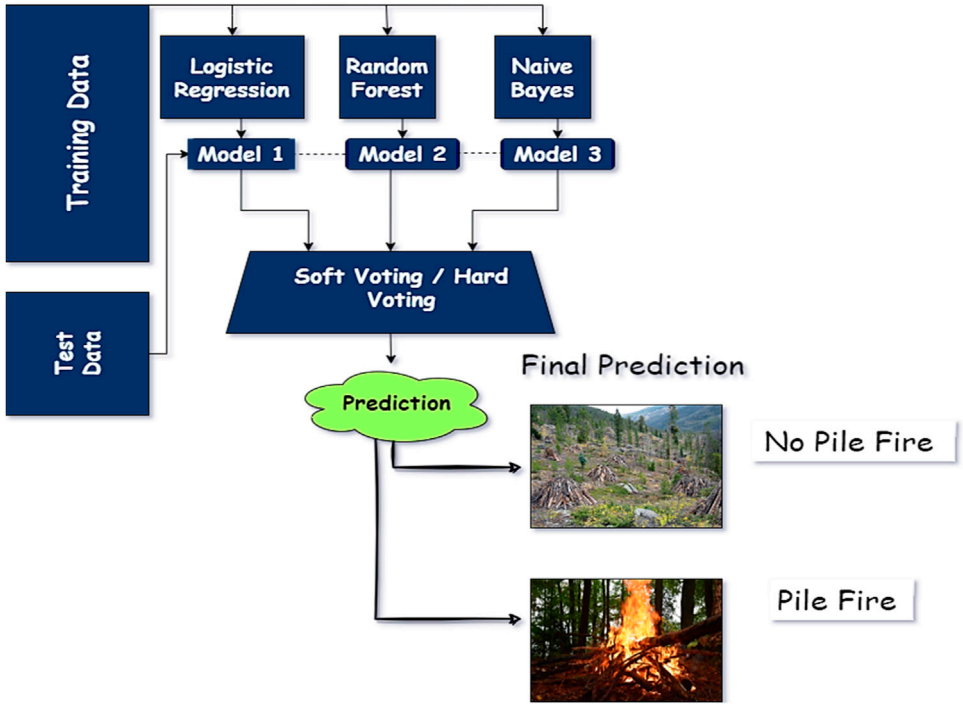


FIGURE 2
Voting Classifier for pile fire classification.

FIGURE 3
Randomly generated frames of pile burn and no pile burn areas using three different drones having normal spectrum.

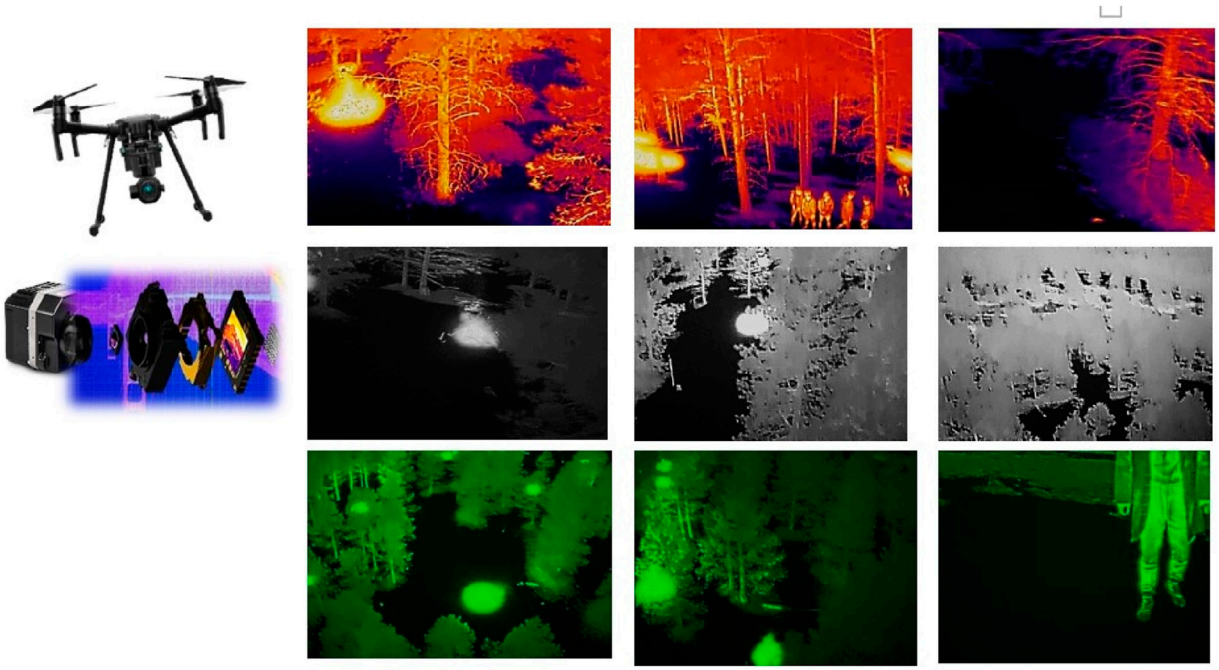


FIGURE 4
Images (frames) of a thermal heat map, comprising thermal-fusion, white-black-fusion, and green-black-fusion palettes, which were taken from the sky view to the ground view.

structured to facilitate easy extraction of frames for analysis. The FLAME collection has a total of four different recordings, which were captured using different drone cameras, such as normal-range-spectrum with RGB color-space shown in Figure 3 and thermal-heat-map, white-hot-fusion and green-hot-fusion palettes shown in Figure 4.

4.1 Normal-spectrum palette

The Zenmuse-X4S and the Phantom-3 Camera were used to capture the Normal-Spectrum Palette. Many videos and photos with pile-fire and no-pile-fire imagery were accessible in the FLAME repository. Two clips were recorded using the Zenmuse-X4S Camera. One was a 16-min video with a resolution of 1,280 × 720 and a frame rate of 29 frames-per-second (fps). At 1,280 × 720 quality and 29 frames per second, a further 6 min of video was available for pile-fire from the start of the burning. The size of the videos was 1.14 GB and 479 MB, respectively. The video captured from the Phantom-3 Camera was of length 17 min with a frame resolution of 3,840 × 2,160. Figure 3 shows some images (frames) of pile-fire and no-pile-fire clips captured by the drone. These clips can be found in the suggested flame dataset. All videos were shot using different palettes—Normal-Spectrum, Thermal-Fusion, White-Hot, and Green-Hot—each captured with specialized drone cameras. All videos were shot at 640 × 512 resolution and 30 frames per second. There were several videos with varying lengths of fire and no-fire types. Table 1 (a) shows the sample set distribution for normal spectrum palette images.

4.2 Thermal-fusion palette

The Thermal-Fusion Palette was captured using a Forward-Looking-Infrared (FLIR) Vue Pro R camera. A solitary video clip was accessible in the FLAME repository, showcasing frames depicting both scenes with pile fires and scenes without pile fires. The video of length 24 min and has 640 × 512 resolution with 30 frames-per-second (fps). The size of the video clip was 2.83 GB. Table 1 (b) shows the sample set distribution for thermal fusion palette images.

4.3 White-hot palette

The White-Hot palette was recorded using a Forward-Looking-Infrared (FLIR)-Vue-Pro-R-camera. Only one video clip was available on the FLAME repository, which has frames containing both pile-fire and no-pile-fire imagery. The video of length 2 min and has 640 × 512 resolution with 30 frames-per-second (fps). The size of the video was 43.3 MB. Table 1 (c) shows the sample set distribution for white-hot palette images.

4.4 Green-hot palette

The Green-Hot Palette was recorded using Forward-Looking-Infrared (FLIR)-Vue Pro-R-camera. Only one video clip was present on the FLAME repository with frames containing both pile-fire and no-pile-fire imagery. The video of length 5 min and has 640 × 512 resolution with 30 frames-per-second (fps). The size of the video

TABLE 1 Training samples and Testing samples of (a) normal spectrum palette, (b) thermal fusion palette, (c) white-hot palette and (d) green-hot palette.

(a) Normal spectrum palette		
Class Label	Training Sample	Testing Sample
Pile Fire	25,018	6,242
No-Pile Fire	14,357	6,107
Total	43,375	12,353
(b) Thermal Fusion Palette		
Class Label	Training Sample	Testing Sample
Pile Fire	686	100
No-Pile Fire	518	118
Total	1,204	218
(c) Whit-hot palette		
Class Label	Training Sample	Testing Sample
Pile Fire	914	100
No-Pile Fire	392	114
Total	1,306	214
(d) Green-hot palette		
Class Label	Training Sample	Testing Sample
Pile Fire	722	118
No-Pile Fire	522	100
Total	1,244	218

clip was 146 MB. Table 1 (d) shows the sample set distribution for green-hot palette images.

5 Image pre-processing and augmentation

5.1 Image pre-processing

Image processing of flame is an important part of fire safety engineering since pixel-based fire detection systems are increasing rapidly. The border extraction method was one of the most critical phases in fire image processing, serving as a precursor and laying the groundwork for subsequent processing. There were several motives because flame margin identification was crucial. For starters, the flame borders provide a framework for measuring many aspects of flame dynamics, such as structure, volume, orientation, and steadiness. Second, establishing flame borders can reduce the file size processed and screen out unwanted data within the frame, such as noise in the background (Celik, 2010), (Habiboğlu et al., 2011). Border identification, in other words, can preserve significant structural characteristics of the flames while shortening time consumption. Finally, border identification can be used to split a group of flames. This was useful in smothering many flames in boiler tubes with a multi-burner system. Moreover, prompt identification of flame borders can trigger a fire alert and offer firefighters information on the type of fire, combustible substances, the outside of the flame, and so on. For example, the mobility of a monitored flame border can be used to distinguish between actual and spurious fire alarms (Çelik and Demirel, 2009), (Umar et al., 2017), (Binti Zaidi et al., 2015). In this study, the authors have extracted features from an image/frame using pixel value, Gabor, and Sobel filters. Figure 5 depicts a sample image of a thermal fusion palette and flame edge detection phases.

5.2 Image augmentation

In this study, we utilized image augmentation techniques to diversify the training dataset, thereby enhancing the model's exposure to a wider range of image variations. By altering factors such as orientation, brightness, and size randomly, the model is compelled to recognize image subjects across different contexts. Unlike image preprocessing, which is applied to both training and test sets, image augmentation is exclusively implemented on the training data. Various methods such as translation, rotation, lighting adjustments, and color manipulation can be employed for data augmentation, which can be categorized into offline and online stages. Offline augmentation involves expanding the training dataset before training begins, while online augmentation increases the diversity of images seen during training. In this research, offline augmentation was employed to create a dataset for plant seedling categorization. This involved resizing, cropping, and horizontally flipping training photos into three sizes. Additionally, data augmentation was illustrated in Figure 6, depicting augmentation and transformation on thermal fusion frames across different spectrums. Such augmentation is vital for enhancing the system's robustness and accuracy in estimating thermal heat within a pile.

6 Experimental results and analysis

6.1 Traditional machine learning techniques and its comparative analysis

6.1.1 Thermal-fusion palette

The total number of frames in the training portion was 1,204, out of which 686 were of class "pile-fire" and 518 frames of class "pile-no-fire." All the frames were scrambled before dumping into the model. The dataset contained many raw frames, blackout frames, and so on. As a result, image processing was required to extract the features from the frame. A total of three filters were applied to each frame to extract features from it. So, the first feature extracted was the pixel value. Each image has a pixel value representing how bright the image is in a specific region of the image or how much intensity of the color is present in that region. The image is distributed into three channels: red, blue, and green. Thus, each channel is represented by 8 bits, and a total of three channels have range from 0 to 255). The second feature extractor was the Gabor filter; now, a total of four Gabor filters were applied on each frame with different parameters. The parameters were theta value, which ranges

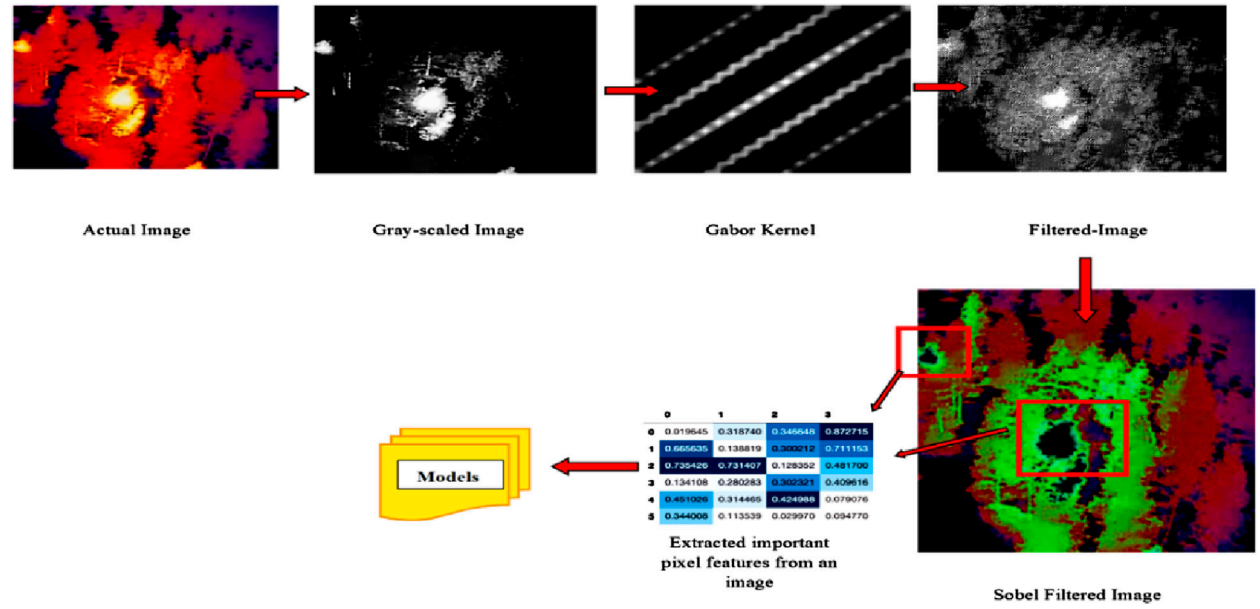


FIGURE 5
Infrared image processing of pile fire images for feature extraction.

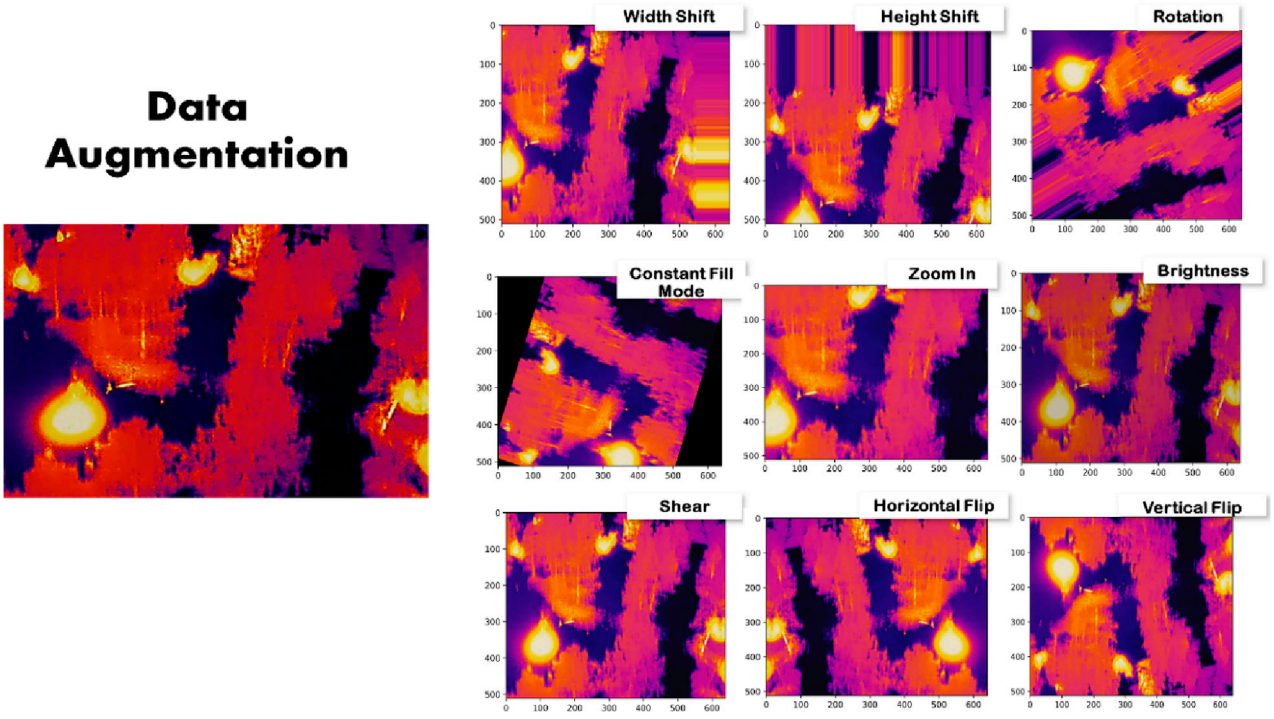


FIGURE 6
Sample augmented images of infrared -thermal processed pile fire images.

from 0.0 to 2.0; sigma value, which ranges between 1 and 3; lambda value of 0.78; and gamma value of 0.5. The third feature extractor was the Sobel filter or the Sobel edge detector, a gradient-based method that looks for significant changes in an image's first derivative. To evaluate the accuracy of the test dataset, 218 frames, including 100 pile-fire-labeled frames and 118 pile-no-fire-labeled frames, were fed into traditional machine learning algorithms such as RF and SVM. Table 2 shows the training parameters used for both algorithms, and Table 3 shows the classification report, volume of the model, inference CPU time,

TABLE 2 Training parameters of (A) RF classifier and (B) SVM classifier.

(A)		(B)	
Parameter	Method	Parameter	Method
N_estimators	100	C	2
Criterion	Gini	Kernel	'rbf'
Max_depth	12	decision_function_shape	'ovo'
Min_samples_split	2	gamma	'auto'
Random-state	42	max_iter	200

and flops. Figures 7A, B show the confusion matrix and ROC-AUC curve for RF algorithm applied on thermal heat map frames to detect fire piles.

6.1.2 White-hot palette

The total number of frames in the training portion was 1,306 frames, out of which 914 frames were of class "pile-fire" and 392 frames of class "pile-no-fire." In image processing for this frame, doing the same thing here by applying three filters on each frame to extract features from the frame. So, the first feature extracted was the pixel value. Each image has a pixel value representing how bright the image is in a specific region of the image or how much intensity of the color is present in that region. The image was distributed into one channel only: Black and White. Thus, each channel is represented by 8 bits, and a total of one channel has 8 bits (ranging from 0 (darker) to 255 (brighter)). The second feature extractor was the Gabor filter; now, four Gabor filters were applied on each frame with different parameters. The parameters were theta value, which ranges from 0.0 to 1.0; sigma value, which ranges between 1 and 4; lambda value of 0.85; and gamma value of 0.7. The third feature extractor was the Sobel filter or the Sobel edge detector, a gradient-based method that looks for significant changes in an image's first derivative. 100 Pile-Fire-labeled frames and 114 Pile-No-fire-labeled frames were used in traditional machine learning methods, such as RF and SVM, to evaluate the accuracy of the test dataset. Table 4 shows the report of training accuracy on the training set and testing accuracy on the testing set. Figures 8A, B show the confusion matrix and ROC-AUC curve for RF algorithm applied on white hot palette frames to detect fire piles.

6.1.3 Green-hot palette

The total number of frames in the training portion was 1,244, of which 722 were of class "pile-fire" and 522 frames of class "pile-no-fire." In image processing for this frame, doing the same thing here by applying three filters on each frame to extract features from the frame. So, the first feature extracted was the pixel value. Each image has a pixel value representing how bright the image is in a specific

TABLE 3 Accuracy for evaluation of the pile fire for thermal fusion classification (224 × 224).

Model name	Size of model (MB)	Dataset type	Training accuracy	Testing accuracy	Precision	Recall	F1-score	AUC
RF	49.67	Thermal- Fusion Palettes	0.997	0.995	0.99	0.99	0.99	0.987
SVM (RBF Kernel)	200	Thermal- Fusion Palettes	0.994	0.981	0.98	0.97	0.98	0.995

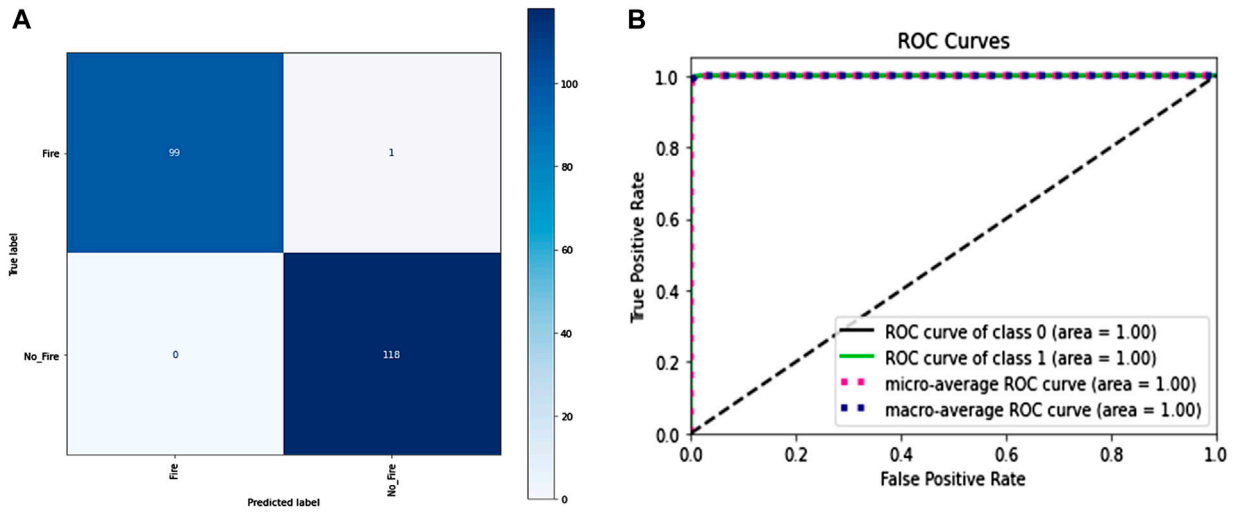


FIGURE 7
(A): Confusion matrix and **(B)** ROC-AUC curve RF algorithm applied on thermal heat map frames to detect fire piles (Best Model Performance).

TABLE 4 Accuracy for evaluation of the pile fire for white-hot palette classification (224 × 224).

Model name	Size of model (MB)	Dataset type	Training accuracy	Testing accuracy	Precision	Recall	F1-score	AUC
RF	49.67	White-Hot Palettes	0.998	0.996	0.99	0.99	0.99	0.996
SVM – RBF Kernel	200	White-Hot Palettes	0.99	0.983	0.99	0.98	0.985	0.997

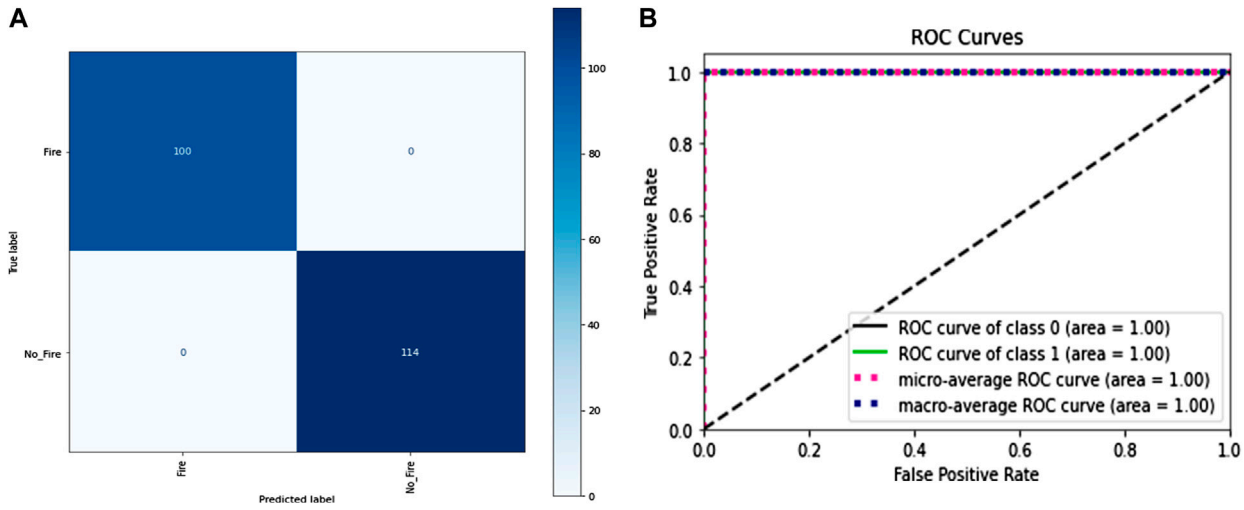


FIGURE 8
(A): Confusion matrix and **(B)** ROC-AUC curve RF algorithm applied on thermal heat map frames to detect fire piles (Best Model Performance).

region of the image or how much intensity of the color is present in that region. The image was distributed into one channel only: Black and Green. But here, the green channel often contains larger bits. This is due to our eyes' increased sensitivity to green color and decreased sensitivity to blue or other colors (ranging from 0 (darker) to 255 (brighter)). The second feature extractor was the Gabor filter. A Gabor filter with different parameters was applied to each frame. The parameters were theta value, which ranges from 0.0 to 2.0; sigma value, which ranges between one and 2; lambda value of 0.78; and gamma value of 0.65. The third feature extractor was the Sobel filter or edge detector, a gradient-based method that looked for significant changes in an image's first derivative. To evaluate the accuracy of the test dataset, 218 frames, including 118 Pile-Fire-labelled frames and 100 Pile-No-fire-labelled frames, were fed into traditional machine learning algorithms such as RF and SVM. Table 5 shows the report of training accuracy on the training set and testing accuracy on the testing set. Figures 9A, B show the confusion matrix and ROC-AUC curve for RF algorithm applied on white hot palette frames to detect fire piles.

6.2 Knowledge transferring technique: Transfer learning (TL)

There are ample advantages to using the RF algorithm. Still, the main drawback of using RF is that it has more decision trees, making the algorithm computationally expensive and complex. It is also not preferable for live streams or real-time analysis. Furthermore, manual feature extraction is difficult to eliminate because it requires business and domain logic to build a robust model that can replicate and capture trends and patterns from data (Çelik et al., 2007), (Khalil et al., 2021), (Chino et al., 2015), (Verstockt et al., 2013). Transfer learning was, therefore, increasingly employed to save cost and time by removing the need to train several machine learning models from the start to execute identical tasks. This paper uses a transfer learning approach to reduce computational costs. The initial layers, except the last dense layer, were frozen, and pre-trained models with trainable parameters contributing to the last layer were used, which reduced the time required for the training and compared to training the entire model from scratch. Pre-trained models were used to compensate for the lack of labeled training data held by an organization. Table 6 shows the training parameters used to implement transfer learning approaches in this study.

6.2.1 Thermal-fusion palettes

The total number of frames in the training portion was 1,204, of which 686 were of class "pile-fire" and 518 frames of class "pile-no-fire." All the frames were scrambled before being sent to the network. Furthermore, data augmentation was performed using different techniques before feeding the training samples into the model. The initial learning rate was set to 0.0001, and the Adamax optimizer was used to regulate weight and minimize the loss

TABLE 5 Accuracy for evaluation of the pile fire for green hot palettes classification (224 × 22).

Model name	Size of model (MB)	Dataset type	Training accuracy	Testing accuracy	Precision	Recall	F1-score	AUC
RF	49.67	Green-Hot Palettes	0.998	0.958	0.96	0.96	0.96	0.997
SVM – RBF Kernel	200	Green-Hot Palettes	0.994	0.912	0.91	0.92	0.91	0.992

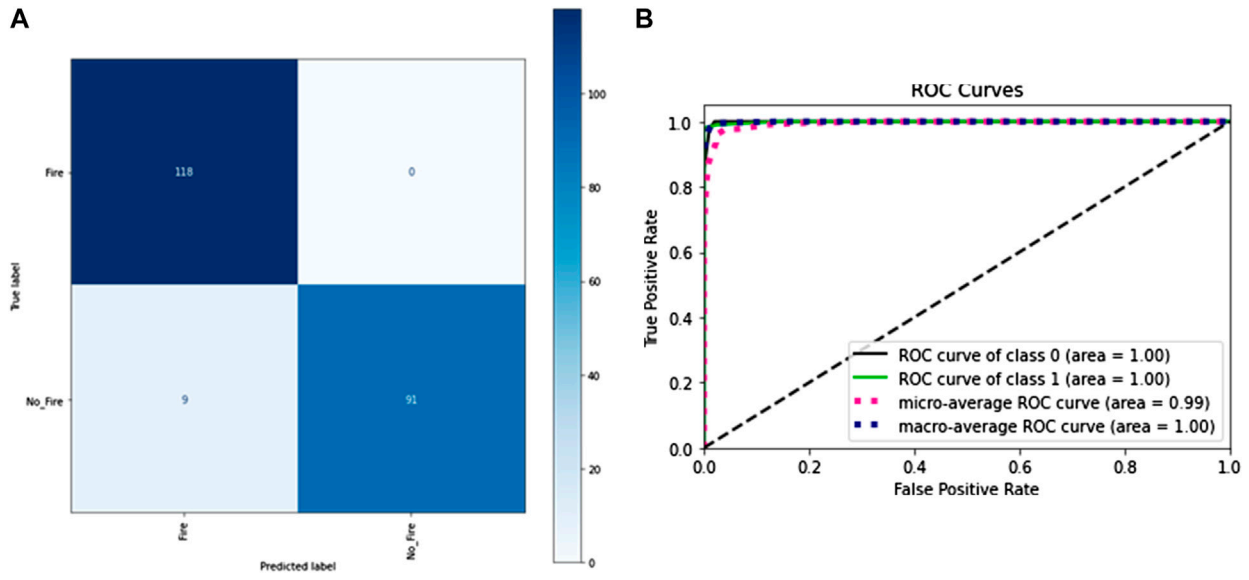


FIGURE 9
(A): Confusion matrix and **(B)** ROC-AUC curve RF algorithm applied on thermal heat map frames to detect fire piles (Best Model Performance).

TABLE 6 Training parameters of transfer learning neural networks.

Parameters	Method
Epoch	5–50
Iteration per epoch	25
Steps per epoch	15
Weight decay	0.00004
Learning rate	0.001 > 0.0005 > 0.00025 > 0.000125
Learning rate decay	0.00001 (Iteration equal step size)
Loss Function	binary_crossentropy
Activation Function	Sigmoid
Optimizer	Adamax
Batch size	50

function. In addition, a batch size of 50 was employed while training the model. To test the performance and error loss of the testing samples, 218 frames were fed into the pre-trained networks, in which 100 frames were labeled as "pile-fire" and 118 frames as "pile-no-fire." Five models were fine-tuned and trained on the Thermal-Fusion Palette dataset. Among them, Resnet50V2 performs better on test and train datasets of the thermal-fusion palette. Table 7 shows the error-loss and accuracies of multiple models. The classification "pile-fire versus no-pile-fire" obtained 99.08% testing accuracy. Figures 10A, B depict training versus validation loss and accuracy curves for Resnet50V2 model.

6.2.2 White-hot palettes

The total number of frames in the training portion was 1,306 frames, out of which 914 frames were of class "pile-fire" and 392 frames of class "pile-no-fire." To test the performance and loss of the testing samples, 214 frames were fed into the pre-trained networks, in which 100 frames were labeled as "pile-fire" and 114 frames as "pile-no-fire." Five models were fine-tuned and trained on the White-Hot Palette dataset. Among them, InceptionV3 performs well on test and train datasets of the white-hot palette. Table 8 shows the training and testing accuracy, model weight, and model inference time on CPU taken while training the model with Bflops and AUC score. In inceptionv3, the researcher got a testing accuracy of around 99.86% for pile fire classification. Figures 11A, B depict training versus validation loss and accuracy curves for the InceptionV3 model.

TABLE 7 Transfer learning enhances the performance of pile-fire classification for thermal fusion. (227 × 227).

Model name	Size of model (MB)	Inference time (ms)	Training accuracy	Testing accuracy	Precision	Recall	F1-score	AUC	BFlops
AlexNet	233	539	0.8512	0.8303	0.9175	0.7542	0.8278	0.8086	0.0122
Inception ResnetV2	92	42.2	0.9538	0.9700	0.9478	1.00	0.9732	0.9727	0.1322
InceptionV3	215	130.2	0.9728	0.9817	0.9914	0.9746	0.9829	0.9991	0.0567
VGG16	528	69.5	0.8945	0.9037	0.9145	0.9068	0.9106	0.9350	0.3069
Resnet50V2	98	45.6	0.9845	0.9908	0.9915	0.9915	0.9915	0.9946	0.0696

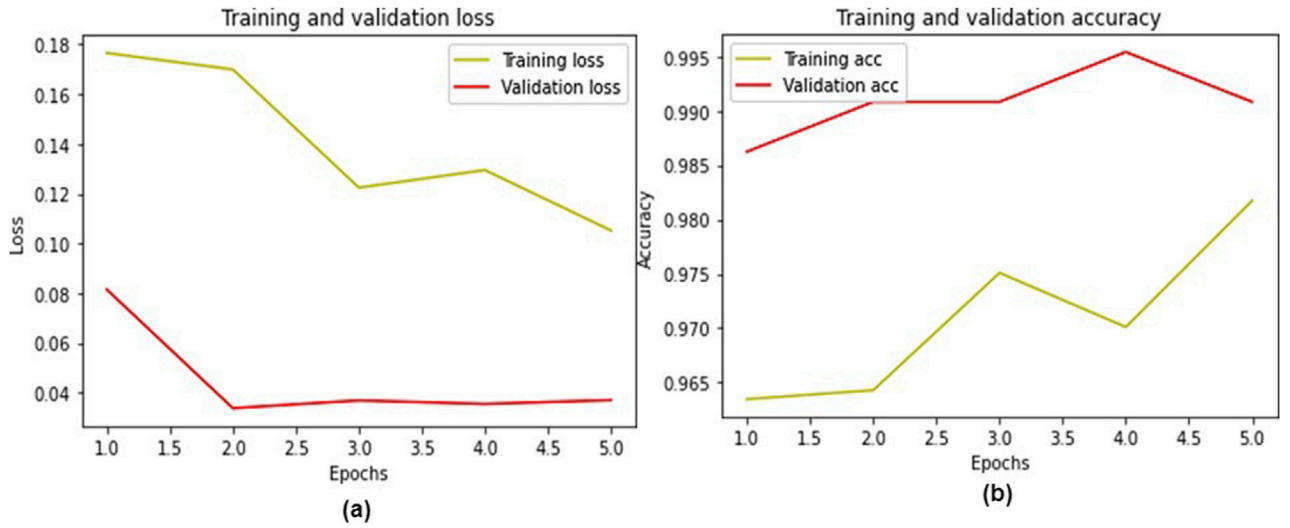


FIGURE 10
(A) Training loss vs. validation loss **(B)** training accuracy vs. validation accuracy for Resnet50V2 (best model performance).

TABLE 8 Transfer learning accuracy for evaluating pile fire classification for white hot palette (227 × 227).

Model name	Size of model (MB)	Inference time (ms)	Training accuracy	Testing accuracy	Precision	Recall	F1-score	Auc-curve	BFlops
AlexNet	233	539	0.9478	0.9019	1.00	0.8158	0.8985	0.9578	0.0122
Inception ResnetV2	92	42.2	0.9850	0.9950	1.00	0.9903	0.9945	1.00	0.1322
InceptionV3	215	130.2	0.9990	0.9986	1.00	1.00	1.00	1.00	0.0567

6.2.3 Green-hot palettes

The total number of frames in the training portion was 1,244, of which 722 were of class "pile-fire" and 522 frames of class "pile-no-fire." To test the performance and loss of the testing samples, 218 frames were fed into the pre-trained networks, in which 118 frames were labeled as "pile-fire" and 100 frames as "pile-no-fire." Five models were fine-tuned and trained on the Green-Hot Palette dataset. Among them, InceptionV3 performed better on a test set of the green-hot palettes. Table 9 shows the accuracy, inference CPU time, and flops while training and testing the model. In the classification of "pile-fire versus no-pile-fire," the researcher obtained a testing accuracy of 99.89%. Figures 12A, B depict training versus validation loss and accuracy curves for the InceptionV3 model.

6.2.4 Normal-spectrum palettes

The total number of frames in the training portion was 43,375, of which 25,018 were of class "pile-fire" and 14,357 were class "pile-no-fire." To test the performance and loss of the testing samples, 12,343 frames were fed into the pre-trained networks, comprising 6,242 frames labeled as "pile-fire" and other 6,107 frames labeled as "pile-no-fire.

Thus, three pre-trained models were fine-tuned and trained on a normal-spectrum palette dataset. Among them, VGG16 performs well on test and train datasets of a normal-spectrum palette. The frames (data) collected from the Matrice 200 drone and Zenmuse X4S camera were used to train the model, while the frames collected from the Phantom drone

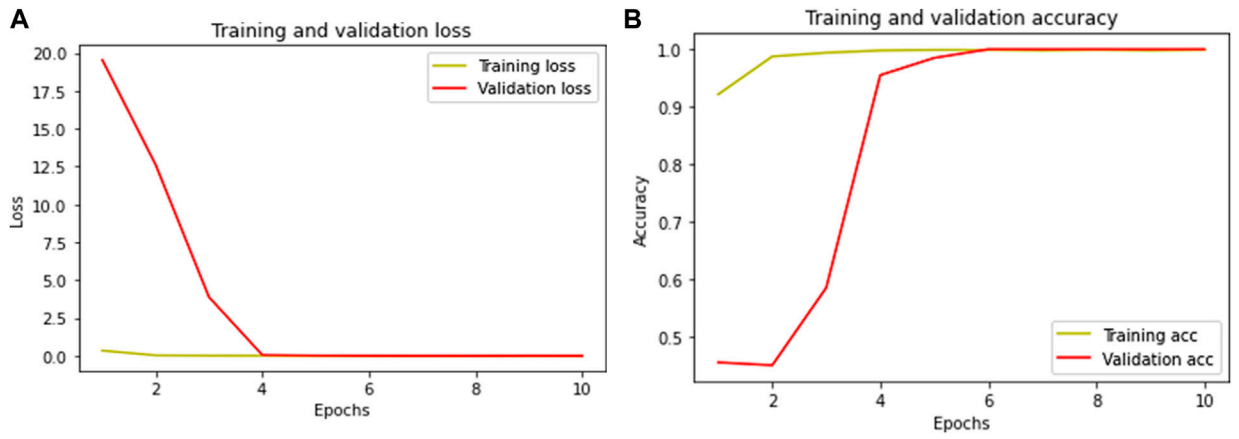


FIGURE 11
(A) Training Loss Vs. Validation Loss **(B)** Training Accuracy for InceptionV3 model (Best Model Performance).

TABLE 9 Transfer learning accuracy for evaluating pile fire classification for a green hot palette (227 × 227).

Model name	Size of model (MB)	Infere nce time (ms)	Training accuracy	Testing accuracy	Precision	Recall	F1-score	Auc-curve	BFlops
AlexNet	233	539	0.9622	0.9587	0.9473	0.9636	0.9553	0.9933	0.0122
Inception ResnetV2	92	42.2	0.9688	0.9400	0.9645	0.9622	0.9633	0.9826	0.1322
InceptionV3	215	130.2	0.9930	0.9989	1.00	1.00	1.00	1.00	0.0567
VGG16	528	69.5	0.9487	0.9850	0.9787	0.9891	0.9838	0.9983	0.3069
Resnet50 V2	98	45.6	0.9396	0.9550	0.900	0.900	0.9473	1.00	0.0696

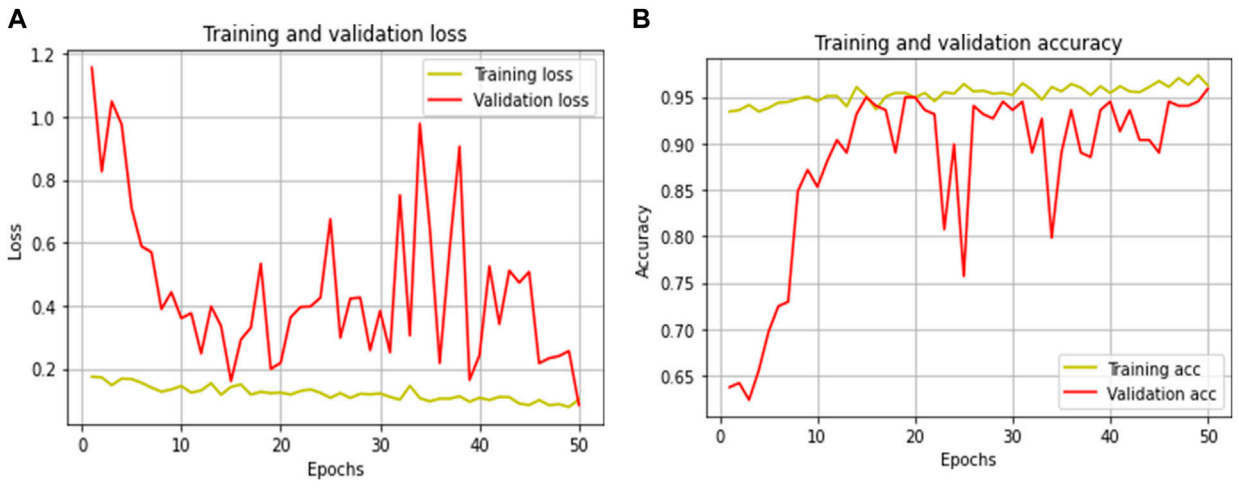


FIGURE 12
(A) Training Loss Vs. Validation Loss **(B)** Training Accuracy for InceptionV3 model (Best Model Performance).

were used to test the model. As a result, there is no overlap between the training and validation samples. Table 10 shows the accuracy, inference CPU time, and flops while training and testing the model. The "pile-fire vs. no-pile-fire" classification attained a testing accuracy of 77%. Figures 13A, B depict the error loss and accuracy curve graphs generated during training and testing for VGG-16. Figure 13C shows the confusion matrix of the VGG-16 model for detecting fire piles.

TABLE 10 Transfer learning accuracy for evaluating pile fire classification for normal spectrum palette (227 × 227).

Model name	Size of model (MB)	Inference time (ms)	Training accuracy	Testing accuracy	Precision	Recall	F1-score	AUC	BFlops
Inception V3	215	130.2	0.9792	0.7525	0.8117	0.5040	0.6218	0.7152	0.0567
VGG16	528	69.5	0.9364	0.7700	0.7186	0.7075	0.7130	0.7641	0.3069
Resnet50 V2	98	45.6	0.9312	0.7276	0.7811	0.4523	1.00	0.7320	0.0696

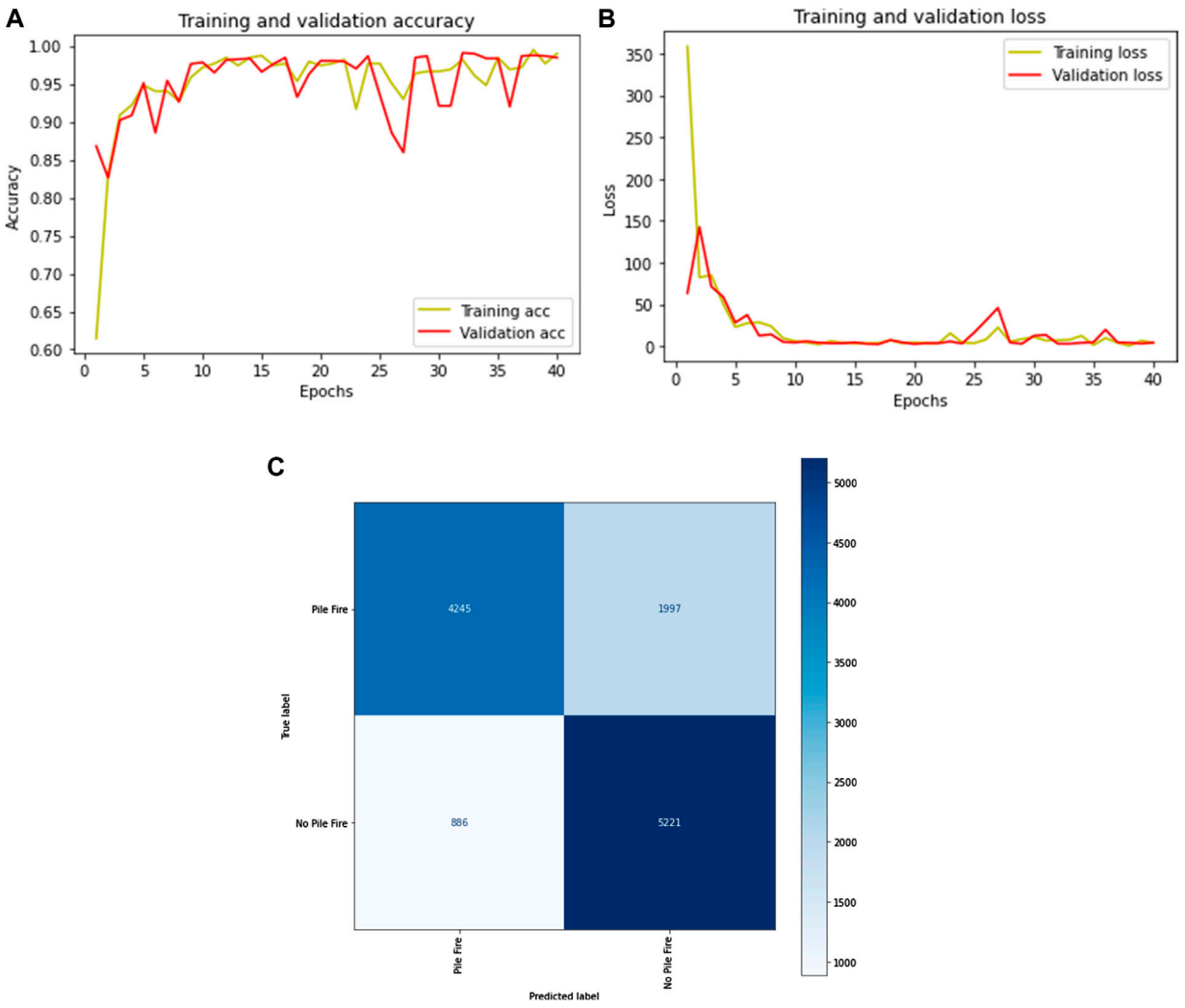


FIGURE 13
(A) Training Loss Vs. Validation Loss **(B)** Training Accuracy for VGG-16 model **(C)** Confusion-matrix of VGG16 model for normal spectrum palette (Best Model Performance).

6.2.5 Optimization of normal–Spectrum model: Transfer learning–Using pre-trained knowledge of VGG- 16 for feature extraction and amalgamation with ensemble learning algorithms

VGG-16 has a deep convolutional layered architecture and uses a smaller kernel size, which reduces the trainable and non-trainable parameters. The input layer for this model was ninety single-dimensional feature vectors. In this study, the authors used a VGG-16 with a single-dimensional kernel space in every convolution layer, followed by a max-pooling layer to extract boundary points of flame features from them. The ninety input vectors as a one-batch and the single-dimensional VGG-16 can grab boundary pixel values as feature vectors with the same weights in each network. Table 11 shows the performance accuracies obtained by amalgamating VGG-16 (feature-extractor) with ensemble learning models. In this study, authors have used many ensemble learning models for the amalgamation process with VGG16 and XGBoost because they selected the optimum tree model. Proposed

TABLE 11 Table of accuracy for evaluation of the pile fire classification of normal spectrum frame using VGG-16 (feature-extractor) + ensemble.

Model name	Size of model (MB)	Training accuracy	Testing accuracy
(Feature-Extractor) VGG16 + XgBoost	528 MB + 490 MB	0.9992	0.75
(Feature-Extractor) VGG16 + Random Forest Classifier	528 MB + 49.67 MB	0.9997	0.82
(Feature-Extractor) VGG16 + AdaBoost (Base Estimator – Logistic Regression)	528 MB + 333 MB	0.9943	0.77
(Feature-Extractor) VGG16 + Gradient Boost	528 MB + 85.8 MB	0.9999	0.77
(Feature-Extractor) VGG16 + Voting Classifiers (Logistic Regression, RF, Naïve Bayes)	600+ MB	0.90 (Hard)/(Soft)	0.91 (Hard)/(Soft)

classifiers used in voting were Logistic Regression, RF, and Naïve Bayes. These types of classification algorithms can be beneficial for balancing the unique shortcomings of a group of equally highly performing models.

6.3 Discussion

This study aimed to assess the effectiveness of fine-tuning five pre-trained convolutional architectures for classifying frames into two classes: "Pile Fire" and "No-Pile-Fire". The primary contribution of this work lies in the systematic evaluation and identification of the most suitable deep transfer learning models for each type of thermal imagery, leading to highly accurate fire detection across diverse environmental conditions. Following are the key contributions.

- Model-specific optimization: The study does a deep analysis of the performance of different CNN models applied to various types of thermal imagery. Of the rest, ResNet50V2 reached 99.08% on thermal fusion frames, AlexNet performed at 95.87% on green-hot frames, InceptionV3 reached 99.87% on white-hot frames, while VGG16 did best with 77.37% on the normal spectrum frame type. These results emphasize the importance of appropriate architecture selection based on the characteristics of the data.
- VGG16 as Feature Extractor: One of the novelties of this study is to use VGG16 with one-dimensional kernel space in each convolution layer, which could obtain boundary pixel values of flame features effectively. This work shows that the ongoing approach blended with max-pooling can be more informative for fine-grained features required for accurate fire detection.
- Ensemble Learning Integration: t also tends to investigate the combination of VGG16 with ensemble learning algorithms such as XGBoost, Logistic Regression, Random Forest, and Naïve Bayes. Such combinations tend to improve predictive performance by cancelling out specific drawbacks of single classifiers and thereby yielding a fire-detection system that is robust in performance and based on the strengths of various algorithms.

7 Conclusion

This study underscores the importance of model selection and optimization in developing reliable fire detection systems, providing a foundation for future research in enhancing disaster management and environmental monitoring technologies. The FLAME dataset includes thermal imagery processed by various filters to extract information from an infrared image. In this study, two different approaches were used for fire classifications using different image modalities. In the first task, a standard machine learning approach, such as binary pile fire classification, was used to label data and grasp data insights. In the second method, researchers presented a transfer learning strategy to execute the sharing of knowledge concepts, which can result in a far more precise and efficient model. In this study, for thermal fusion frames, ResNet50V2 exhibited the highest accuracy of 99.08%; for green hot frames, InceptionV3 achieved the highest accuracy of more than 99%; for white-hot frames, InceptionV3 exhibited the highest accuracy of 99.87%; and lastly, for normal spectrum frames, VGG16 achieved the highest accuracy of 77.37%. Finally, to improve the performance of the normal-spectrum vision model, we combined transfer learning with ensemble learning to produce the most efficient model with an accuracy of 91%. The scientific community can enhance the given findings by building more advanced systems. Another potential application for this collection is the development of fire detection and segmentation algorithms, such as semantic segmentation or instance segmentation, based on a collective study of several imaging modalities, including conventional and thermal images provided by the researcher. Instance segmentation can be used to segregate pile burns, thermal heat, and forest, animals, and human beings so that we can save them before any disaster occurs in that location. Researchers may also use fire segmentation approaches to identify associated networking and monitoring challenges, such as optimum job scheduling for a fleet of drones to cover pile burns in a certain location in the least amount of time.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://ieee-dataport.org/open-access/flame-dataset-aerial-imagery-pile-burn-detection-using-drones-uavs.

Author contributions

DJ: Writing–original draft, Data curation, Methodology, Software, Validation. SaK: Conceptualization, Supervision,

Writing–review and editing. SP: Conceptualization, Supervision, Writing–review and editing. PK: Conceptualization, Supervision, Writing–review and editing. ShK: Supervision, Writing–review and editing. KK: Conceptualization, Supervision, Writing–review and editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Research Support Fund (RSF) of Symbiosis International (Deemed University), Pune, India.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdusalomov, A. B., Islam, B. M. S., Nasimov, R., Mukhiddinov, M., and Whangbo, T. K. (2023). An improved forest fire detection method based on the Detectron2 model and a deep learning approach. *Sensors* 23 (3), 1512. doi:10.3390/s23031512

Afghah, F., Razi, A., Chakareski, J., and Ashdown, J. (2019). "Wildfire monitoring in remote areas using autonomous unmanned aerial vehicles," in *Infocom 2019 - IEEE conference on computer communications workshops, INFOCOM WKSHPS 2019*.

Afghah, F., Shamsoshoara, A., Njilla, L., and Kamhoua, C. (2018a). "A reputation-based stackelberg game model to enhance secrecy rate in spectrum leasing to selfish IoT devices," in *Infocom 2018 - IEEE conference on computer communications workshops*.

Afghah, F., Zaeri-Amirani, M., Razi, A., Chakareski, J., and Bentley, E. (2018b). "A coalition formation approach to coordinated task allocation in heterogeneous UAV networks," in *Proceedings of the American control conference*.

Aggarwal, R., Soderlund, A., Kumar, M., and Grymin, D. (2020). "Risk aware SUAS path planning in an unstructured wildfire environment," in *Proceedings of the American control conference*.

Binti Zaidi, N. I., Binti Lokman, N. A. A., Bin Daud, M. R., Achmad, H., and Chia, K. A. (2015). Fire recognition using RGB and YCbCr color space. *ARPN J. Eng. Appl. Sci.* 10 (21), 9786–9790. Available at: https://www.arpnjournals.org/jeas/research_papers/rp_2015/jeas_1115_2983.pdf

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., et al. Towards federated learning at scale: system design. 2019.

Borges, P. V. K., and Izquierdo, E. (2010). A probabilistic approach for vision-based fire detection in videos. *IEEE Trans. Circuits Syst. Video Technol.* 20 (5), 721–731. doi:10.1109/tcsvt.2010.2045813

Brik, B., Ksentini, A., and Bouaziz, M. (2020). Federated learning for UAVs-enabled wireless networks: use cases, challenges, and open problems. *IEEE Access* 8, 53841–53849. doi:10.1109/access.2020.2981430

Casbeer, D. W., Kingston, D. B., Beard, R. W., and Mc lain, T. W. (2006). Cooperative forest fire surveillance using a team of small unmanned air vehicles. *Int. J. Syst. Sci.* 37 (6), 351–360. doi:10.1080/00207720500438480

Celik, T. (2010). Fast and efficient method for fire detection using image processing. *ETRI J.* 32 (6), 881–890. doi:10.4218/etrij.10.0109.0695

Çelik, T., and Demirel, H. (2009). Fire detection in video sequences using a generic color model. *Fire Saf. J.* 44 (2), 147–158. doi:10.1016/j.firesaf.2008.05.005

Çelik, T., Özkaramanli, H., and Demirel, H. (2007). "Fire and smoke detection without sensors: image processing based approach," in *European signal processing conference*.

Chino, D. Y. T., Avalhais, L. P. S., Rodrigues, J. F., and Traina, A. J. M. (2015). "BoWFire: detection of fire in still images by integrating pixel color and texture analysis," in *Brazilian symposium of computer graphic and image processing*.

Edge TPU - run inference at the edge | Google cloud. Available at: https://cloud.google.com/edge-tpu

Frizzi, S., Kaabi, R., Bouchouicha, M., Ginoux, J. M., Moreau, E., and Fnaiech, F. (2016). Convolutional neural network for video fire and smoke detection. *IECON 2016 - 42nd Annu. Conf. IEEE Industrial Electron. Soc.*, 877–882. Available at: https://api.semanticscholar.org/CorpusID:3541656. doi:10.1109/IECON.2016.7793196

Habiboğlu, Y. H., Günay, O., and Çetin, A. E. (2011). "Flame detection method in video using covariance descriptors," in *ICASSP, IEEE international conference on acoustics, speech and signal processing - proceedings*.

Habiboğlu, Y. H., Günay, O., and Çetin, A. E. (2012). Covariance matrix-based fire and flame detection method in video. *Mach. Vis. Appl.* 23 (6), 1103–1113. doi:10.1007/s00138-011-0369-1

Hu, C., Tang, P., Jin, W., He, Z., and Li, W. (2018). "Real-time fire detection based on deep convolutional long-recurrent networks and optical flow method," in *Chinese control conference, CCC*.

Jadon, A., Omama, M., Varshney, A., Ansari, M. S., and Sharma, R. (2019), FireNet: a specialized lightweight fire and smoke detection model for real-time IoT applications.

Khalil, A., Rahman, S. U., Alam, F., Ahmad, I., and Khalil, I. (2021). Fire detection using multi color space and background modeling. *Fire Technol.* 57 (3), 1221–1239. doi:10.1007/s10694-020-01030-9

Kundu, S., Mahor, V., and Gupta, R. (2018). "A highly accurate fire detection method using discriminate method," in *2018 international conference on advances in computing, communications and informatics, ICACCI 2018*.

Mueller, M., Karasev, P., Kolesov, I., and Tannenbaum, A. (2013). Optical flow estimation for flame detection in videos. *IEEE Trans. Image Process.* 22 (7), 2786–2797. doi:10.1109/tip.2013.2258353

Muhammad, K., Ahmad, J., Lv, Z., Bellavista, P., Yang, P., and Baik, S. W. (2019). Efficient deep CNN-based fire detection and localization in video surveillance applications. *IEEE Trans. Syst. Man. Cybern. Syst.* 49 (7), 1419–1434. doi:10.1109/tsmc.2018.2830099

National Fire News National interagency fire center. Available at: https://www.nifc.gov/fire-information/nfn.

NVIDIA Jetson Nano developer Kit | NVIDIA developer Available at: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-nano/product-development/

Qi, X., and Ebert, J. (2009). A computer vision-based method for fire detection in color videos. *Int. J. Imaging* 2 (9 S). Available at: https://www.usu.edu/cs/people/XiaojunQi/Promotion/IJI.FireDetection.09.pdf

Qureshi, W. S., Ekpanyapong, M., Dailey, M. N., Rinsurongkawong, S., Malenichev, A., and Krasotkina, O. (2016). QuickBlaze: early fire detection using a combined video processing approach. *Fire Technol.* 52 (5), 1293–1317. doi:10.1007/s10694-015-0489-7

Rafiee, A., Dianat, R., Jamshidi, M., Tavakoli, R., and Abbaspour, S. (2011). "Fire and smoke detection using wavelet analysis and disorder characteristics," in *ICCRD2011 - 2011 3rd international conference on computer research and development*.

Rinsurongkawong, S., Ekpanyapong, M., and Dailey, M. N. (2012). "Fire detection for early fire alarm based on optical flow video processing," in *2012 9th international conference on electrical engineering/electronics, computer, telecommunications and information technology, ECTI-CON 2012*.

Saydirasulovich, S. N., Mukhiddinov, M., Djuraev, O., Abdusalomov, A., and Cho, Y. I. (2023). An improved wildfire smoke detection based on YOLOv8 and UAV images. *Sensors (Basel)* 23 (20), 8374. doi:10.3390/s23208374

Shafiq, M., and Gu, Z. (2022). Deep residual learning for image recognition: a survey. *Appl. Sci. Switz.* 12, 8972. doi:10.3390/app12188972

Shamsoshoara, A. Ring oscillator and its application as physical unclonable function (PUF) for password management. 2019.

Shamsoshoara, A., and GitHub repository, GitHub. 2020b. Fire-detection-uav-aerial-image-classification-segmentationunmannedaerialvehicle.

Shamsoshoara, A., Afghah, F., Razi, A., Mousavi, S., Ashdown, J., and Turk, K. (2020a). An autonomous spectrum management scheme for unmanned aerial vehicle networks in disaster relief operations. *IEEE Access* 8, 58064–58079. doi:10.1109/access.2020.2982932

Shamsoshoara, A., Afghah, F., Razi, A., Zheng, L., Fulé, P. Z., and Blasch, E. (2021). Aerial imagery pile burn detection using deep learning: the FLAME dataset. *Comput. Netw.* 193, 108001. doi:10.1016/j.comnet.2021.108001

Statistics National interagency fire center. Available at: https://www.nifc.gov/fire-information/statistics.

Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *31st AAAI conference on artificial intelligence, AAAI 2017*.

Umar, M. M., De Silva, L. C., Bakar, M. S. A., and Petra, M. I. (2017). State of the art of smoke and fire detection using image processing. *Int. J. Signal Imaging Syst. Eng.* 10 (1–2), 22. doi:10.1504/ijsise.2017.084566

Verstockt, S., Beji, T., De Potter, P., Van Hoecke, S., Sette, B., Merci, B., et al. (2013). Video driven fire spread forecasting (f) using multi-modal LWIR and visual flame and smoke data. *Pattern Recognit. Lett.* 34 (1), 62–69. doi:10.1016/j.patrec.2012.07.018

Wang, T., Shi, L., Yuan, P., Bu, L., Hou, X., and Jiang, C. (2017). HIF1α-Induced glycolysis metabolism is essential to the activation of inflammatory macrophages. *Mediat. Inflamm. Proc. - 2017 Chin. Autom. Congr.*, 2017, 9029327, doi:10.1155/2017/9029327

Wu, H., Li, H., Shamsoshoara, A., Razi, A., and Afghah, F. (2020). "Transfer learning for wildfire identification in UAV imagery," in *2020 54th annual conference on information Sciences and systems, CISS 2020*.

Xuan Truong, T., and Kim, J. M. (2012). Fire flame detection in video sequences using multi-stage pattern recognition techniques. *Eng. Appl. Artif. Intell.* 25 (7), 1365–1372. doi:10.1016/j.engappai.2012.05.007

Ye, D. H., Zikic, D., Glocker, B., Criminisi, A., and Konukoglu, E. (2013). Modality propagation: coherent synthesis of subject-specific scans with data-driven regularization. *5MB MODEL SIZE. ICLR17.* 16 (Pt 1), 606–613. doi:10.1007/978-3-642-40811-3_76

Yuan, C., Liu, Z., and Zhang, Y. (2015b). "UAV-based forest fire detection and tracking using image processing techniques," in *2015 international conference on unmanned aircraft systems, ICUAS 2015*.

Yuan, C., Liu, Z., and Zhang, Y. (2017). Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. *J. Intelligent Robotic Syst. Theory Appl.* 88 (2–4), 635–654. doi:10.1007/s10846-016-0464-7

Yuan, C., Zhang, Y., and Liu, Z. (2015a). A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Can. J. For. Res.* 45 (7), 783–792. doi:10.1139/cjfr-2014-0347

Zhang, Q., Xu, J., Xu, L., and Guo, H. Deep convolutional neural networks for forest fire detection. In 2016.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016). "Learning deep features for discriminative localization," in *Proceedings of the IEEE computer society conference on computer vision and pattern recognition*.

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 25 September 2024
DOI 10.3389/fenvs.2024.1408370

OPEN ACCESS

EDITED BY
Sayali Sandbhor,
Symbiosis International University, India

REVIEWED BY
Zhenkun Liu,
Nanjing University of Posts and Telecommunications, China
Bu Zhao,
University of Michigan, United States
Rushikesh Kulkarni,
Symbiosis International University, India

*CORRESPONDENCE
Yong Wang,
wy@zjut.edu.cn

RECEIVED 28 March 2024
ACCEPTED 12 September 2024
PUBLISHED 25 September 2024

CITATION
Wang Y, Tian S and Zhang P (2024) Novel spatio-temporal attention causal convolutional neural network for multi-site PM2.5 prediction. *Front. Environ. Sci.* 12:1408370. doi: 10.3389/fenvs.2024.1408370

Novel spatio-temporal attention causal convolutional neural network for multi-site PM2.5 prediction

Yong Wang*, Shuang Tian and Panxing Zhang

School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China

Multi-site PM2.5 prediction has emerged as a crucial approach, given that the accuracy of prediction models based solely on data from a single monitoring station may be constrained. However, existing multi-site PM2.5 prediction methods predominantly rely on recurrent networks for extracting temporal dependencies and overlook the domain knowledge related to air quality pollutant dispersion. This study aims to explore whether a superior prediction architecture exists that not only approximates the prediction performance of recurrent networks through feedforward networks but also integrates domain knowledge of PM2.5. Consequently, we propose a novel spatio-temporal attention causal convolutional neural network (Causal-STAN) architecture for predicting PM2.5 concentrations at multiple sites in the Yangtze River Delta region of China. Causal-STAN comprises two components: a multi-site spatio-temporal feature integration module, which identifies temporal local correlation trends and spatial correlations in the spatio-temporal data, and extracts inter-site PM2.5 concentrations from the directional residual block to delineate directional features of PM2.5 concentration dispersion between sites; and a temporal causal attention convolutional network that captures the internal correlation information and long-term dependencies in the time series. Causal-STAN was evaluated using one-year data from 247 sites in mainland China. Compared to six state-of-the-art baseline models, Causal-STAN achieves optimal performance in 6-hour future predictions, surpassing the recurrent network model and reducing the prediction error by 8%–10%.

KEYWORDS

air quality prediction, convolutional neural network, PM2.5 dispersion, domain knowledge, feedforward network, spatio-temporal analysis

1 Introduction

Air pollution has emerged as a significant challenge to environmental sustainability in the 21st century. Specifically, PM2.5 increasingly impacts urban health and the quality of life of residents negatively (Yan et al., 2022). Chronic exposure to PM2.5 may elevate the risk of non-communicable respiratory diseases, cardiovascular diseases, and diabetes (Yang et al., 2022). Additionally, short-term exposure to PM2.5 has been shown to accelerate aging, as evidenced by changes in DNA methylation profiles associated with blood coagulation, oxidative stress, and systemic inflammation (Gao et al., 2022). Consequently, PM2.5 prediction studies hold significant importance and are considered a critical issue for environmental protection (Ai et al., 2019). However, current methods that predict

PM2.5 concentrations using data from a single monitoring station fall short in capturing spatial correlations among stations, thereby limiting their predictive accuracy. Therefore, the development of a multi-site PM2.5 prediction approach that harnesses spatial relationships between air quality monitoring stations is crucial. Such an approach would not only broaden the scope of PM2.5 concentration predictions but also enhance the accuracy of these models, providing stronger support for effective air quality management strategies and improving public health protection (Wang et al., 2023).

Over the past few decades, PM2.5 prediction models have evolved from traditional physical and statistical models to more sophisticated machine learning and deep learning approaches. Traditional physical models have primarily been focused on simulating the dispersion, deposition, and chemical reactions of PM2.5 processes (Marvin et al., 2022). For instance, the WRF/Chem model forecasts environmental conditions by assessing the potential physicochemical impacts and dynamics of pollutants (Grell et al., 2005). However, these models' reliance on complex data structures and limited generalization ability present significant challenges for practical applications (Liu and Chen, 2020). Common statistical models employed in PM2.5 prediction include multiple linear regression (MLR) (Lagesse et al., 2020), autoregressive integrated moving average (ARIMA) (Wang et al., 2017), and autoregressive conditional heteroskedasticity (ARCH) (Wu and Kuo, 2012). A primary limitation of these models is their reliance on linear assumptions, which often fails to capture the inherently nonlinear behavior of PM2.5, thus compromising their predictive accuracy (Marsha and Larkin, 2019; Erden, 2023). In contrast, machine learning models excel in capturing nonlinear patterns, thus proving highly effective in PM2.5 prediction. These models harness complex data patterns and relationships, enhancing prediction accuracy by addressing the nonlinearity of PM2.5 concentrations (Wang et al., 2024b). Notable examples include support vector machines (SVM) (Lai et al., 2021), extreme gradient boosting (XGBoost) (Liu et al., 2021), random forest (Chen et al., 2023), and various integration techniques (Sun et al., 2023; Teng et al., 2023; Liu et al., 2024). However, despite their capabilities, many machine learning approaches primarily focus on single-station predictions and often neglect the spatial interactions and distributions among multiple monitoring stations, which limits their effectiveness in comprehensive urban air quality management (de Hoogh et al., 2018).

With advancements in deep learning, researchers are increasingly exploring multi-site PM2.5 forecasting nationwide. Convolutional Neural Networks (CNNs) are pivotal in extracting spatial correlation features from time series, initially recognizing spatial correlations between adjacent values, then leveraging convolutional operators to augment learning processes (Faraji et al., 2022). Recurrent Neural Networks (RNNs), known for sequence modeling, utilize a hidden state vector updated sequentially with each input, facilitating temporal information transmission across time steps (Young et al., 2018). These attributes make RNNs suitable for time series prediction tasks, including PM2.5 forecasting (Shakya et al., 2023). Given that multi-site PM2.5 prediction requires mastering spatio-temporal sequences, hybrid models combining CNNs and RNNs have become prevalent. These models employ CNNs to delineate spatial correlations between stations, while RNNs handle the temporal dynamics of PM2.5 concentrations at individual stations (Chiang and Horng, 2021; Du et al., 2021; Zhang et al., 2022; Teng et al., 2023). However, despite their advantages, RNNs encounter significant challenges in managing long series and large-scale multi-site forecasting due to their limitations in parallel processing and capturing distant dependencies, as highlighted by Liang and Tang (2022), along with others (Vaswani et al., 2017; Khandelwal et al., 2018). To address these limitations, recent studies have focused on developing novel feedforward models that better accommodate the complexities of multi-site PM2.5 prediction (Chinatamby and Jewaratnam, 2023).

Recent advances in multi-site PM2.5 prediction have increasingly focused on innovative feedforward models, notably those employing Temporal Convolutional Networks (TCNs) and attention mechanisms. The significant advantage of TCNs in time series prediction is attributed to their straightforward structure, extensive expansion flexibility, and clear causal constraints (Bi et al., 2022; Li et al., 2022; Nasr Azadani and Boukerche, 2022). For example, Zhang et al. (2021) designed a causal convolutional neural network for short-term PM2.5 prediction using TCNs, whose convolution operation explicitly takes causality into account, i.e., the output of a time step only depends on the same or earlier time steps in the previous layer, providing a new perspective on PM2.5 feedforward prediction modeling. However, TCNs struggles to capture the dependency of distant locations in the time series and to extract the internal correlation information of the input data. The Airformer, recently introduced by Liang et al. (2023), stands as a notable model founded on the attention mechanism for air quality prediction. It depends entirely on this mechanism to discern the spatio-temporal patterns of air quality data and employs the Generation Model and the Inference Model to grasp the inherent uncertainty within the air quality data. While attention-based models demonstrate substantial predictive capabilities, their large size poses practical limitations. Furthermore, the Temporal Convolutional Attention-based Network (TCAN), developed by Hao et al. (2020) for natural language processing tasks, shows promise in sequential task modeling by integrating TCNs with attention mechanisms. However, its application to multi-site PM2.5 prediction is potentially limited by insufficient consideration of spatial correlations among monitoring sites and a lack of domain-specific knowledge on PM2.5 dispersion.

To address the limitations of existing feedforward networks in the task of multi-site PM2.5 prediction, we aim to develop a novel architecture based on feedforward techniques that aspires to reach predictive performance comparable to recurrent networks while improving simplicity and efficiency. Furthermore, the accuracy of multi-site PM2.5 predictions not only depends on advanced modeling techniques but also requires a deep understanding of domain-specific knowledge about the pollutant, such as the direction of pollutant dispersion during events (Wang et al., 2020; Zhou et al., 2021). It is crucial to integrate this knowledge into the prediction models to improve accuracy.

Based on these objectives, this study introduces a novel exploratory architecture: the spatio-temporal attention causal convolutional neural network (Causal-STAN) for multi-site PM2.5 prediction. First, we propose a multi-site spatio-temporal

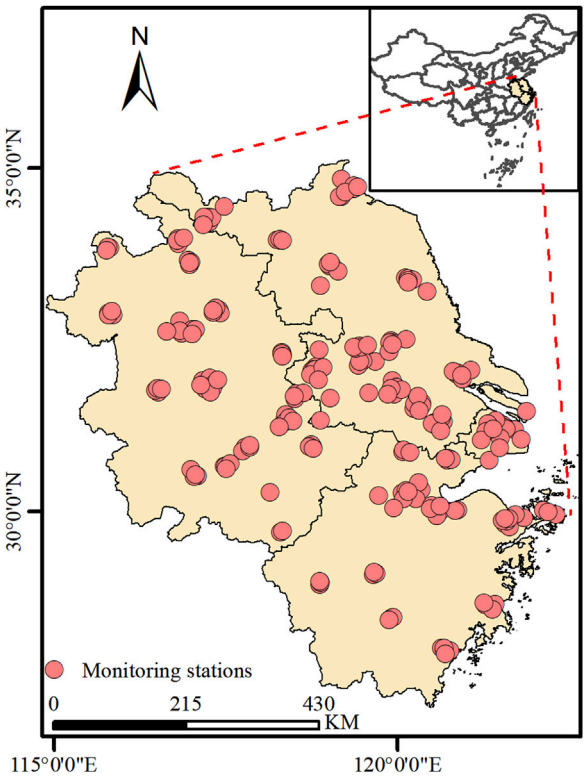


FIGURE 1
Geographical distribution of the 247 air quality monitoring stations across 41 cities in the Yangtze River Delta region, China.

feature integration module. This module employs a CNN to identify temporal local correlation trends and spatial correlations in spatio-temporal data, integrates domain knowledge of PM2.5 pollutant dispersion, and introduces a directional residual block to extract directional features of PM2.5 concentration dispersion between sites. Second, we designed a temporal causal attention convolutional network, inspired by TCAN, that simulates the input causality of RNNs using dilated causal convolution. This network incorporates an attention mechanism to effectively capture the internal correlation information and long-term dependencies within PM2.5 time series. We aim to achieve an approximate replacement of RNNs with temporal causal attention convolutional network and assess the efficacy of the networks' unique "causal" properties in the PM2.5 prediction task. To assess the proposed method's effectiveness, experiments were conducted with data from 247 air quality monitoring stations across 41 cities in the Yangtze River Delta region of China. The main contributions are as follows.

(1) Addressing the scarcity of concise and efficient feedforward prediction models in PM2.5 prediction, this study introduces a novel spatio-temporal attention causal convolutional neural network (Causal-STAN) tailored for multi-site PM2.5 concentration prediction.
(2) Incorporating domain knowledge of air pollutant dispersion, a directional residual block was designed and integrated into the multi-site spatio-temporal feature integration module, enabling the extraction of directional features of inter-site PM2.5 concentration dispersion.
(3) Maximum information coefficients are employed to simultaneously detect similar sites processed by the proposed model, facilitating the extraction of comprehensive knowledge from the dataset.
(4) Performance evaluation results demonstrate that the proposed multi-site PM2.5 feedforward prediction model offers significant advantages over the baseline model, surpassing even the recurrent models in comparison. This model presents a viable alternative to RNNs for multi-site PM2.5 prediction tasks, showcasing its potential effectiveness.

2 Materials and methods

2.1 Study area and data collection

Data for this study were sourced from 247 air quality monitoring stations, spanning the period 1 January 2022, to 31 December 2022, across the Yangtze River Delta region of China, encompassing 41 cities. Figure 1 illustrates the geographical distribution of all study monitoring stations. Hourly data from each monitoring station were compiled into a one-dimensional eigenvector, incorporating pollutant data, meteorological data, and the air quality index (AQI). Pollutant data and AQI were sourced from the urban air quality real-time publishing platform of the China National Environmental Monitoring Center[1], amounting to 4,327,440 records. Meteorological data were acquired from the National Climatic Data Center[2], totaling 1,078,423 records. Detailed data specifications are provided in Table 1. Since the meteorological stations and air pollutant monitoring stations are not directly matched, we selected meteorological data from the nearest meteorological station (based on Euclidean distance) for each pollutant monitoring station. This approach ensures that the meteorological data closely reflect local environmental conditions.

2.2 Data preprocessing

Air quality data collected by the urban air quality real-time publishing platform of the China National Environmental Monitoring Center exhibit a certain percentage of missing values and outliers. Missing values and outliers can result from prolonged operation of monitoring equipment or exposure to extreme weather conditions, such as heavy rain, storms, and haze. Table 1 presents the rates of missing and outlier values for pollutant data and AQI in the dataset, indicating a higher incidence of missing values for AQI. In this study, outliers were detected using the interquartile range (IQR) method, defined as the difference between the third quartile (Q3) and the first quartile (Q1). The upper outlier limit is calculated as Q3 plus 1.5 times the IQR, and the lower limit is Q1 minus 1.5 times the IQR. Data identified as outliers are subsequently

1 https://air.cnemc.cn:18007/

2 https://www.ncei.noaa.gov/

TABLE 1 Description of dataset in this study, the last two columns show the missing and outlier rates for the pollutant indicators in the dataset, respectively.

Data	Feature	Unit	Miss rate	Outlier rate
AQIs	AQI	-	4.8%	1.8%
Pollutants	PM2.5	μg/m^3	2.6%	2.1%
	PM10	μg/m^3	2.0%	1.6%
	SO_2	μg/m^3	1.7%	1.2%
	NO_2	μg/m^3	1.8%	1.6%
	O_3	μg/m^3	2.0%	0.7%
	CO	mg/m^3	1.9%	1.2%
Meteorology	Temperature	℃	-	-
	Atmospheric Pressure	mmHg	-	-
	Humidity	%	-	-
	Wind Direction	Encoding	-	-
	Wind Speed	m/s	-	-
	Visibility	Km	-	-
	Precipitation	mm	-	-

replaced with missing values. To address the missing values in the dataset, two-way linear interpolation was employed for imputation.

2.3 Spatio-temporal correlation analysis

Considering the evident spatio-temporal correlation of PM2.5 across various monitoring stations, constructing a model based solely on historical data from a single station may present limitations. To enhance the accuracy of model predictions, it is imperative to integrate data from multiple monitoring stations for spatio-temporal correlation analysis. For a dataset with historical data from N monitoring stations, input data are transformed into a three-dimensional tensor $\boldsymbol{X} \in \boldsymbol{R}^{N\times T\times F}$, where T denotes the time step and F denotes the number of features associated with PM2.5. For any given target site, the initial step involves determining the correlation between its time series data and that of other sites.

The maximum information coefficient (MIC) is a nonparametric statistical method that measures the correlation between two variables, as proposed by Reshef et al. (2011). The MIC is defined as shown in Equations 1, 2:

$$\boldsymbol{I}(\boldsymbol{x},\boldsymbol{y}) = \int \boldsymbol{p}(\boldsymbol{x},\boldsymbol{y}) log_2 \frac{\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})}{\boldsymbol{p}(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{y})} \boldsymbol{dxdy} \tag{1}$$

$$\boldsymbol{MIC}(\boldsymbol{x},\boldsymbol{y}) = \underset{\boldsymbol{a}\times\boldsymbol{b}<\boldsymbol{B}}{\boldsymbol{max}} \frac{\boldsymbol{I}(\boldsymbol{x},\boldsymbol{y})}{\boldsymbol{log_2}\, \boldsymbol{min}(\boldsymbol{a},\boldsymbol{b})} \tag{2}$$

where x and y are two random variables, and a and b are the number of bins into which the x and y-axes are divided, respectively. B is a parameter whose size is approximately the 0.6 power of the sample size (Zhu et al., 2021). MIC values range from 0 to 1, where 0 indicates no correlation and 1 indicates a perfect correlation. Unlike traditional methods such as the Pearson correlation coefficient, the MIC is capable of capturing more complex nonlinear relationships and offers greater robustness (Wang et al., 2024a). We employ MIC to quantify interactions between sites and effectively capture nonlinear relationships between variables, as shown in Equation 3:

$$\boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_i) = \boldsymbol{MIC}(\boldsymbol{Y}_\star, \boldsymbol{Y}_i) \tag{3}$$

where $\boldsymbol{Y}_\star$ and $\boldsymbol{Y}_i$ represent the PM2.5 time series of the current target site and the ith site among all study sites, respectively; $\boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_i)$ the numerical value of correlation.

Upon obtaining the correlation values between the current target site and all other sites, these values are compiled into a correlation vector, as shown in Equation 4:

$$\boldsymbol{\theta}_\star = [\boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_1), \cdots, \boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_i), \cdots, \boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_N)] \tag{4}$$

Considering that not all sites significantly impact the target site, setting a correlation threshold to filter out sites with strong interactions with the target site is prudent. As a result, the spatio-temporal correlation analysis produces the final input feature vector as follows, as shown in Equation 5:

$$\boldsymbol{X}_\star = \{\boldsymbol{X}_i | \boldsymbol{\theta}(\boldsymbol{Y}_\star, \boldsymbol{Y}_i) > \boldsymbol{\theta}_{th}, \boldsymbol{i} \in \{\boldsymbol{1}, \boldsymbol{2}, \cdots, \boldsymbol{N}\}\} \tag{5}$$

where $\boldsymbol{\theta}_{th}$ is the set correlation threshold, and $\boldsymbol{X}_\star \in \boldsymbol{R}^{C\times T\times F}$ is the input feature vector for the target site after filtering based on spatio-temporal information. In this vector, C represents the number of monitoring stations that are significantly correlated with the target site, as determined by exceeding the correlation threshold $\boldsymbol{\theta}_{th}$.

2.4 Architecture of the proposed network

The architecture of the proposed Causal-STAN is illustrated in Figure 2, which comprises two main components. The first component is the multi-site spatio-temporal feature integration

10.3389/fenvs.2024.1408370

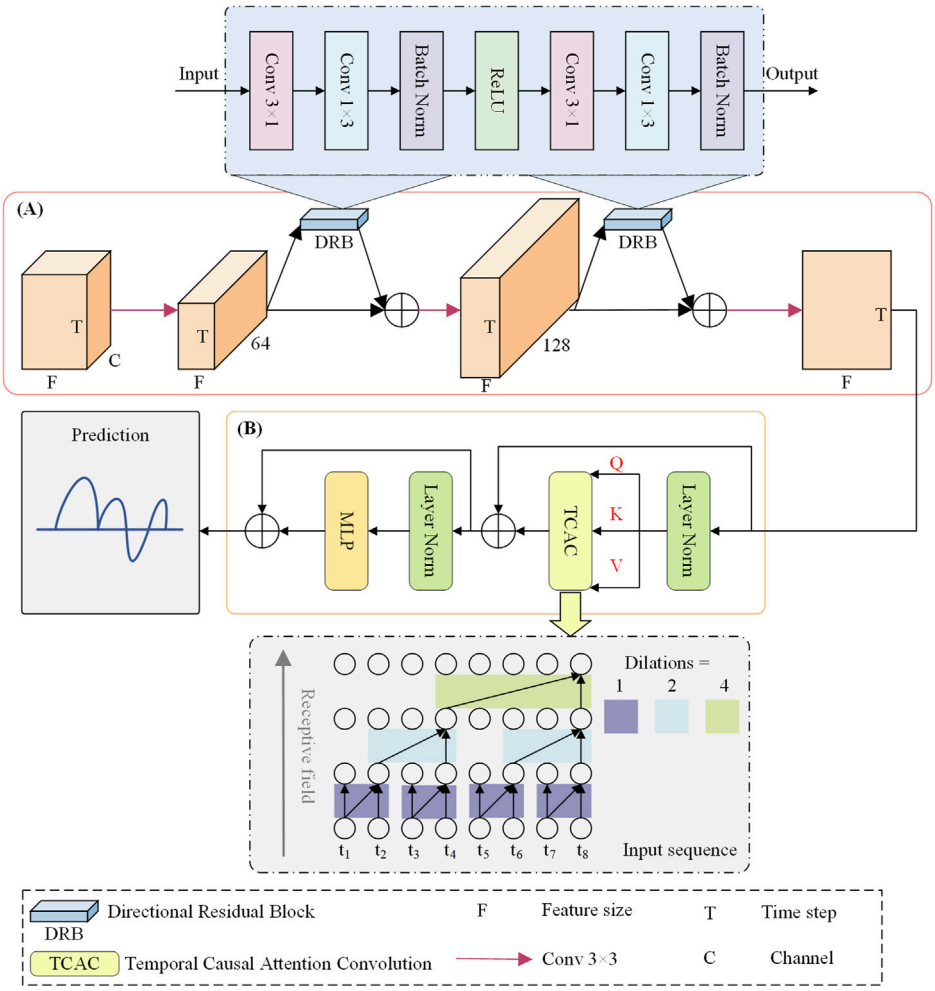


FIGURE 2
The proposed framework of Causal-STAN, which consists of two components. **(A)** Multi-site spatio-temporal feature integration module; **(B)** temporal causal attention convolutional network, which we illustrate using eight time-step examples, here, kernel size = 2, dilation = 1, 2, 4.

module, depicted in Figure 2A, while the second component is the temporal causal attention convolutional network, depicted in Figure 2B. The following subsections elaborate on the details of these two components.

2.4.1 Multi-site spatio-temporal feature integration module

Given the input feature $\boldsymbol{X}_{\star}$, a 3-dimensional vector with spatio-temporal data from various sites, we propose a multi-site spatio-temporal feature integration module to extract spatio-temporal features across sites. The design of this module incorporates the following considerations: 1) **Spatio-temporal dependence:** Acknowledging the spatial dimension's significance in spatio-temporal data for feature extraction, we integrate a feature dimension transformation strategy within the convolution operation to enhance the model's capability to extract spatio-temporal dependent information from various monitoring stations. 2) **Directional features of PM2.5 concentration diffusion:** With domain knowledge of air pollution diffusion, recognizing that pollutants typically exhibit directional diffusion among stations, affected by environmental factors like wind direction, we designed a directional residual block to capture these directional features.

In Figure 2A, as an example, the first feature at the initial time step has channels corresponding to the number of monitoring sites in $\boldsymbol{X}_{\star}$. Different target sites may have varying numbers of related sites, denoted by the parameter C, which plays a key role in the prediction model. The number of input channels in the first convolutional layer is set to match the C value of the current target site. This allows the model to dynamically adjust the input channels for each site, optimizing the use of information from related sites. The initial 3 × 3 convolution is used to identify temporal local correlation trends and spatial correlations in the spatio-temporal data, expanding the feature dimension from the specific *C* value to 64. In other words, the initial 3 × 3 convolution expands the feature dimension of the target site's feature vector from the specific *C* value to 64. Subsequently, the directional residual block is introduced with this feature dimension, utilizing longitudinal Conv 3 × 1 and horizontal Conv 1 × 3 convolution kernels to pinpoint pollutants' spatial spreading characteristics on different axes, thereby improving the model's detection of potential spatial correlations among sites. Following this, a second 3 × 3 convolution further

enhances the extraction of temporally localized correlation trends and spatial correlations, increasing the spatio-temporal data's feature dimensions to 128. At this juncture, a second directional residual block is applied to the higher-dimensional spatio-temporal data, aiming to learn deeper directional features between sites. Ultimately, a 3 × 3 convolution completes the feature integration, reducing the spatio-temporal data's feature dimension to 1 and generating a two-dimensional synthesized feature vector as input for the temporal causal attention convolutional network.

The directional residual block uses two convolution operations: a longitudinal 3 × 1 kernel and a horizontal 1 × 3 kernel. These kernels are designed to capture spatial patterns along specific axes, enhancing the model's ability to analyze pollutant diffusion across the dataset.

2.4.1.1 Longitudinal convolution (3 × 1)

This kernel is structured to extend vertically over three rows in a single column and is configured to extract correlations across different features or sites, analyzing changes along a vertical axis within the data. Such an arrangement is pivotal for identifying site-specific pollution trends or environmental factors that may consistently influence adjacent sites. By targeting vertical slices of data, this convolution adeptly captures dependencies arising from vertical stratification of atmospheric components or variations in emission sources among proximately located sites.

2.4.1.2 Horizontal convolution (1 × 3)

The horizontal kernel, spreading over one row and three columns, is engineered to monitor temporal sequences, enabling the model to trace the evolution of environmental conditions or pollutant levels over time. This kernel excels at detecting patterns across three consecutive time steps, offering insights into how dynamic environmental conditions, such as shifts in wind direction or speed, impact the dispersion and concentration of pollutants.

The outputs from these directional convolutions are integrated using a residual learning framework, where a skip connection adds the block's input to its output. This method is instrumental in mitigating the vanishing gradient problem commonly encountered in deep neural networks, while also preserving identity information throughout the layers. Such an approach enables the model to refine its predictions by continuously learning from the discrepancies between predicted and actual pollution patterns, significantly enhancing the model's accuracy and sensitivity to subtle environmental variations.

2.4.2 Temporal causal attention convolutional network

Beyond spatial dependence, a site's air quality is influenced by its historical data. Given a site's hidden state $\boldsymbol{I}_{1:T}^{(L)} \in \boldsymbol{R}^{T\times F}$ (output from the spatio-temporal feature integration module) across all past steps, the basic feedforward neural network TCN fails to learn the temporal dependence of distant locations within the sequence and cannot extract the internal correlation information of the inputs. Here, we introduce the temporal causal attention convolutional network as an effective alternative to the standard TCN for temporal modeling. As shown in Figure 2B, the key part of the proposed network is the temporal causal attention convolution, featuring two main modifications:

2.4.2.1 Temporal causal attention

To address TCN's limitation, we introduce a self-attention mechanism that captures internal time-series relationships and long-range dependencies. Unlike standard attention, this mechanism preserves the causal order of events.

2.4.2.2 Dilated causal convolution

Since air quality at the current time step is only influenced by past events, we maintain the causal structure by applying dilated causal convolution, which ensures the model respects the correct temporal order.

The computation of intermediate variables within the temporal causal attention convolution is illustrated in Figure 3:

1. Temporal causal attention is applied, as shown in Equation 6:

$$\boldsymbol{IA}_{1:T}^{(L)} = \boldsymbol{TCA}\left(\boldsymbol{I}_{1:T}^{(L)}\right) \tag{6}$$

where $\boldsymbol{IA}_{1:T}^{(L)}$ represents the intermediate representation after applying temporal causal attention on the first T time steps.

2. Given $\boldsymbol{IA}_{1:T}^{(L)}$, we apply dilated causal convolution to it, as shown in Equation 7:

$$\boldsymbol{I}_{1:T}^{(L+1)} = \boldsymbol{DCC}\left(\boldsymbol{IA}_{1:T}^{(L)}\right) \tag{7}$$

where $\boldsymbol{I}_{1:T}^{(L+1)}$ represents the output of the dilated causal convolution, as illustrated in the top square on the right side of Figure 3. To maintain the output tensor's length equal to the input tensor's, zero padding of length $((\boldsymbol{k}-\boldsymbol{1})\boldsymbol{2}^{\boldsymbol{l}-\boldsymbol{1}})$ is added to the input tensor's left side, as depicted in the Zero Padding on the left side of Figure 3, ensuring causal convolution.

3. The full temporal causal attention convolutional network is built by stacking L layers of temporal causal attention convolutions, covering both depth and time.

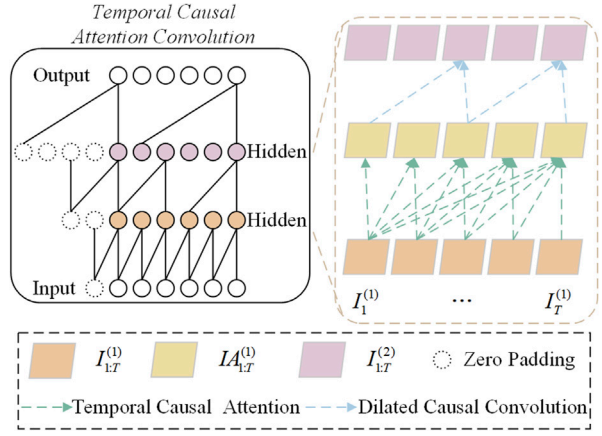


FIGURE 3
Temporal causal attention convolution, right-hand side, is illustrated using an example of five time steps, here, kernel size = 2 and dilation = 2.

2.4.3 Temporal causal attention

Temporal causal attention, illustrated in Figure 4, involves integrating the influences of previous time steps into the current time step. Distinct from the self-attention structure, our attention mechanism selectively uses information from previous time steps only, thanks to a masking mechanism and optimized weight matrix, ensuring relevance to the current and preceding time steps while blocking future interactions.

Initially, three linear transformations, $f(\mathrm{T})$, $g(\mathrm{T})$, and $h(\mathrm{T})$, map $\boldsymbol{I}_{1:T}^{(L)}$ to $\boldsymbol{d_k}$- dimensional $\boldsymbol{queries}(\boldsymbol{q}_{1:T}^{L} = \boldsymbol{f}(\boldsymbol{I}_{1:T}^{(L)}))$, $\boldsymbol{keys}(\boldsymbol{k}_{1:T}^{L} = \boldsymbol{g}(\boldsymbol{I}_{1:T}^{(L)}))$, and $\boldsymbol{d_v}$- dimensional $\boldsymbol{values}(\boldsymbol{v}_{1:T}^{L} = \boldsymbol{h}(\boldsymbol{I}_{1:T}^{(L)}))$. Subsequently, batch matrix multiplication calculates the dot products between $\boldsymbol{k}_{1:T}^{L}$ and $\boldsymbol{q}_{1:T}^{L}$, dividing each result by $\sqrt{\boldsymbol{d_k}}$ to compute the weight matrix $\boldsymbol{WA}^{(L)}$, as shown in Equation 8:

$$\boldsymbol{W}_{i,j}^{(L)} = \frac{\boldsymbol{k}_i^{(L)T} \cdot \boldsymbol{q}_j^{(L)}}{\sqrt{\boldsymbol{d_k}}} \tag{8}$$

Next, an upper triangular masking matrix $\boldsymbol{M}^{(L)}$ is constructed, with elements on and below the diagonal set to 0, and those above the diagonal set to 1. The masking matrix $\boldsymbol{M}^{(L)}$ is then applied to $\boldsymbol{W}^{(L)}$, as shown in Equation 9:

$$\boldsymbol{Wi}_{i,j}^{(L)} = \begin{cases} \boldsymbol{W}_{i,j}^{(L)}, \ \boldsymbol{if}\ \boldsymbol{M}_{i,j}^{(L)} = \boldsymbol{0} \\ -\infty, \ \boldsymbol{if}\ \boldsymbol{M}_{i,j}^{(L)} = \boldsymbol{1} \end{cases} \tag{9}$$

For $\boldsymbol{i}, \boldsymbol{j} = \boldsymbol{1}, \boldsymbol{2}, \cdots, \boldsymbol{T}$. Finally, a softmax function is applied to normalize $\boldsymbol{Wi}^{(L)}$ to obtain $\boldsymbol{WA}^{(L)}$. This effectively suppresses the masked weights in $\boldsymbol{Wi}^{(L)}$ to nearly zero, thus isolating the future time step from the current one. Given $\boldsymbol{v}_i^{(L)}$, the weighted output is calculated using the normalized weights, as shown in Equation 10:

$$\boldsymbol{IA}_t^{(L)} = \sum_{i=0}^{t} \boldsymbol{WA}_i^{(L)} \cdot \boldsymbol{v}_i^{(L)} \tag{10}$$

Here, $\boldsymbol{IA}_t^{(L)}$ represents the output of the temporal attention, serving as the input for dilated causal convolution.

2.4.4 Dilated causal convolution

For a one-dimensional time series input $\boldsymbol{x} \epsilon \boldsymbol{R}^{\boldsymbol{n}}$ and a convolution kernel $\boldsymbol{f}$: $\{\boldsymbol{0}, \cdots, \boldsymbol{k-1}\} \rightarrow \boldsymbol{R}$, the dilated convolution operation F on an element s in the series is defined as shown in Equation 11:

$$\boldsymbol{F}(\boldsymbol{s}) = (\boldsymbol{x} *_{\boldsymbol{d}} \boldsymbol{f})(\boldsymbol{s}) = \sum_{i=0}^{k-1} \boldsymbol{f}(\boldsymbol{i}) \cdot \boldsymbol{x}_{\boldsymbol{s}-\boldsymbol{d}\cdot\boldsymbol{k}} \tag{11}$$

Here, $*$ denotes the standard convolution operation, d the dilation factor, k the size of the convolution kernel, and $\boldsymbol{s} - \boldsymbol{d} \cdot \boldsymbol{k}$ considers the elements in the past direction. When $d = 1$, the dilated convolution simplifies to a regular convolution. The dilated convolution expands the receptive field by increasing the size of the convolution kernel k and the dilation factor d. To encompass long-term historical data, the dilation factor follows an exponential growth strategy by setting d for layer i to $\boldsymbol{2}^{\boldsymbol{i}}$.

Aligned with temporal causal attention, dilated causal convolution preserves causality in sequence prediction, preventing future time information from influencing the model. This causality is essential for the PM2.5 prediction task as it guarantees predictions are based solely on past and present data, excluding future information.

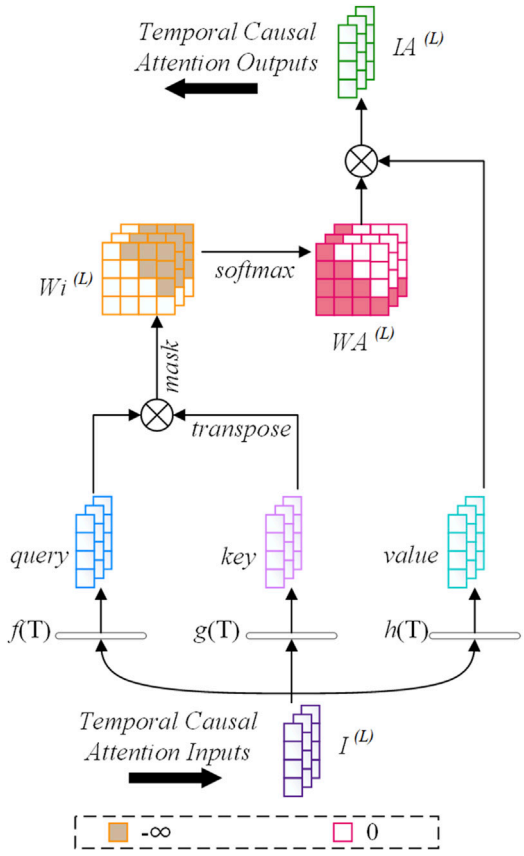


FIGURE 4
Illustration of the temporal causal attention mechanism, showing the integration of previous time steps' influences into the current time step.

2.5 Experimental setup

The entire dataset is divided into three parts: training, validation, and test sets. The training set comprises the first 60% of each month's data, while the validation set contains the last 20% of each month. Finally, the remaining 20% of data in each month is allocated to the test set. This division aims to preserve the time series continuity and adapt to monthly environmental changes, enhancing the model's relevance to real-world scenarios. Our model forecasts 6-hour future predictions based on the past 24-hour readings, setting $T = 24$. To train our model, we employ the Adam optimizer, utilizing MSE as the loss function, with a batch size of 32. Through a grid search spanning the range {2, 3, 4, 5, 6, 7, 8} for the number of layers (levels) and kernel size in the temporal causal attention convolutional network, it was found that optimal performance is achieved with both parameters set to 4. For a more detailed analysis of these hyperparameters and their impact on model performance, please refer to Section 3.4. The model underwent 150 iterations, employing an early stopping strategy to prevent overfitting. Model performance evaluation employed three metrics: root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination ($\boldsymbol{R}^2$), as shown in Equations 12–14:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (12)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (14)$$

Here, n is the total number of samples, y_i the ith observed value, $\hat{y}_i$ the ith predicted value, and $\bar{y}$ the average of observed values. Lower RMSE and MAE values signify higher prediction accuracy, while R^2 ranges between 0 and 1, with values closer to 1 denoting a better model fit.

2.6 Experimental baseline model

Traditional time series prediction models like LSTM and GRU effectively address the issue of vanishing and exploding gradients in recurrent models through the use of gating mechanisms, showcasing robust performance in PM2.5 prediction tasks. These models were included in our comparative experiments.

CNN-BiLSTM (Du et al., 2021), among the earliest deep network models, adopts a joint spatio-temporal prediction approach by integrating CNN with BiLSTM.

ST-CausalConvNet (Zhang et al., 2021) highlights causality's critical role in PM2.5 prediction, inspiring the development of our proposed temporal causal attention convolutional network.

TCAN (Hao et al., 2020), an exploratory feedforward sequence prediction network in NLP, combines convolution with an attention mechanism to approximate recurrent networks, laying the groundwork for our proposed temporal causal attention convolutional network. This model is included in a comparative experiment to underscore the significance of inter-site spatio-temporal feature extraction in multi-site PM2.5 prediction tasks.

DAGCGN (Tariq et al., 2023), developed in 2023 as a distance-adaptive graph convolutional gated recurrent network, excels in identifying complex spatio-temporal interactions between neighboring monitoring sites. Employing GCN in conjunction with the recurrent network GRU for multi-site PM2.5 prediction, this model demonstrates the recurrent model's significant prediction performance and serves as our primary comparison model.

3 Results and discussion

3.1 Spatio-temporal correlation analysis result

We examined the default thresholds for spatio-temporal correlation analysis in the proposed model at four sample monitoring stations. Figure 5 shows the number of correlated stations identified at six different thresholds for sites numbered 1233, 1270, 1155, and 1141. Notably, these four sample sites are located in four distinct provinces within the dataset. As illustrated in Figure 5, we selected six specific thresholds—0.35, 0.40, 0.45, 0.50, 0.55, and 0.60—which are critical values between 0.35 and 0.60. At these thresholds, the number of correlated sites identified varies significantly, potentially impacting the model's performance.

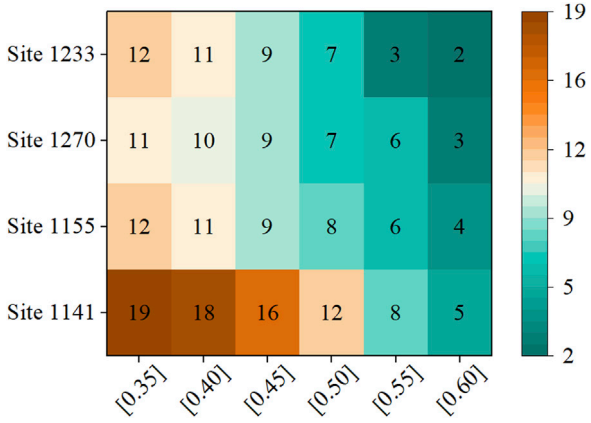


FIGURE 5
The number of relevant sites obtained at different thresholds for the four sample sites, with each row representing the number of relevant sites obtained at six different thresholds for that site.

Therefore, we focused on analyzing the effect of these thresholds on model accuracy. Additionally, we excluded thresholds below 0.35 and above 0.60 because they could either overly complicate or oversimplify the predictive model's input feature vector, unlikely providing an optimal threshold. Specifically, thresholds below 0.35 might introduce a large number of sites with weak or no relevance to the target site, cluttering the model's input feature vector with irrelevant features and reducing prediction accuracy. Conversely, thresholds above 0.60 could significantly reduce the number of correlated sites, potentially degrading the model to a single-site prediction and neglecting spatial correlation features between sites. We further investigated the optimal threshold among the selected values.

Figure 6 shows the RMSE for the four sample sites when predictions are made using Causal-STAN at various thresholds. The results indicate a trend of decreasing and then increasing RMSE values for all four sites, with prediction error minimized at a threshold of 0.50. This suggests that thresholds set too high or

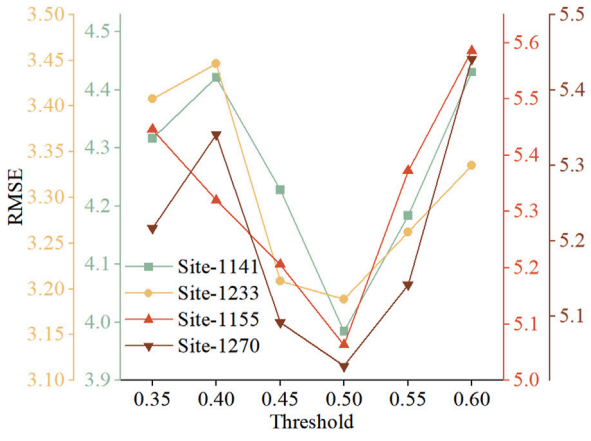


FIGURE 6
RMSE predicted for the four sample sites at different thresholds using Causal-STAN.

too low adversely affect prediction outcomes: a high threshold limits spatially associated sites, losing key information, while a low threshold incorporates weakly related data, increasing noise and potential inaccuracies. Thus, a default threshold of 0.50 was adopted for Causal-STAN's spatio-temporal correlation analysis in further model evaluations.

3.2 Model comparison results

To validate the effectiveness of our model, we conducted a performance evaluation using the test sets from all sites in the dataset. The final forecasting results for each model were determined by averaging the outcomes across 247 target sites. This approach helps to comprehensively assess the overall performance of each model across all sites, providing a thorough understanding of their predictive capabilities. Figure 7 displays a performance comparison between the proposed model and six competing models for forecasting the next 6 h, with results averaged over the 247 target sites. The results indicate that Causal-STAN outperforms the other models in predicting the next 6 h, achieving the lowest RMSE and MAE, along with the highest R^2, whereas the classical GRU method exhibits the weakest performance. The performance among CNN-BiLSTM, ST-CausalConvNet, and DAGCGN is closely matched, with DAGCGN slightly leading; however, none surpass Causal-STAN.

Table 2 provides a comprehensive enumeration of the average results for the three evaluation metrics across each model, offering a thorough numerical comparison that complements the visual representation depicted in Figure 7. Experimental results indicate that TCAN significantly outperforms traditional time-series prediction models such as GRU and LSTM. This superior performance is largely attributed to TCAN's unique model architecture. Unlike traditional recurrent prediction models based on RNN architecture, TCAN employs a TCN combined with an attention mechanism to predict PM2.5 concentrations. TCAN integrates the unique dilated convolutions of TCN with an attention mechanism, effectively expanding its receptive field in the time series and capturing long-term dependencies within the time series data. These features significantly enhance its performance in PM2.5 prediction tasks. However, a notable limitation of TCAN in multi-site PM2.5 prediction is its inability to perform spatial feature extraction. In contrast, CNN-BiLSTM surpasses TCAN by effectively capturing spatial dependencies between sites through CNN, yet it has its limitations. Although BiLSTM considers temporal features bi-

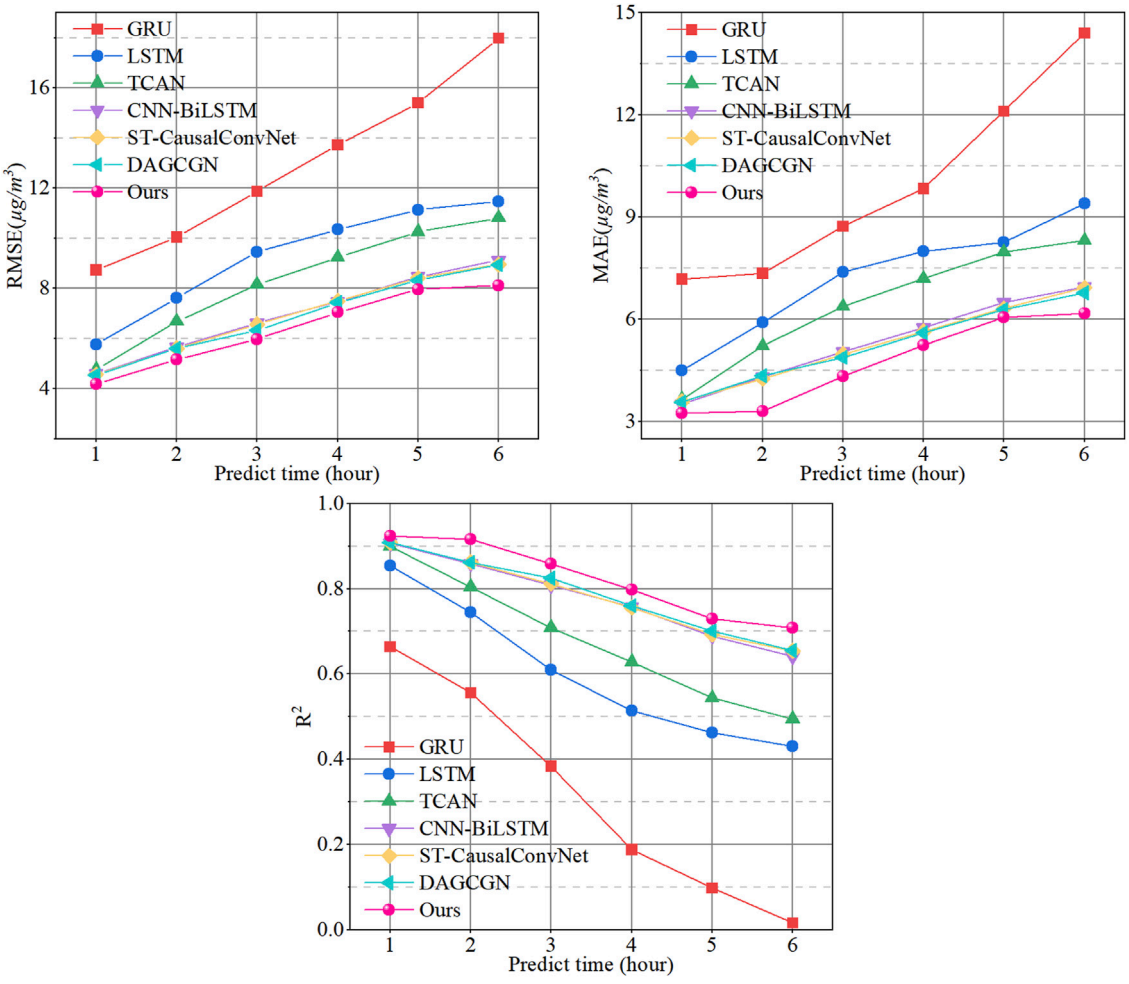


FIGURE 7
Comparison of the proposed Causal-STAN (**Ours**) with six other methods for hourly prediction, utilizing test data from all 247 stations.

TABLE 2 Average metrics for each model based on test data from all 247 stations, with best results highlighted in bold.

Model	Metric	1 h	2 h	3 h	4 h	5 h	6 h
GRU	RMSE	8.735	10.054	11.866	13.721	15.396	17.982
	MAE	7.182	7.354	8.734	9.841	12.108	14.405
	R^2	0.664	0.556	0.384	0.188	0.098	0.016
LSTM	RMSE	5.754	7.629	9.452	10.342	11.130	11.465
	MAE	4.501	5.901	7.383	7.998	8.253	9.397
	R^2	0.854	0.744	0.609	0.514	0.463	0.430
TCAN	RMSE	4.765	6.679	8.152	9.242	10.254	10.813
	MAE	3.655	5.218	6.377	7.199	7.976	8.316
	R^2	0.899	0.804	0.709	0.628	0.544	0.495
CNN-BiLSTM	RMSE	4.588	5.654	6.617	7.464	8.466	9.126
	MAE	3.517	4.324	5.057	5.753	6.503	6.955
	R^2	0.907	0.859	0.808	0.757	0.689	0.641
ST-CausalConvNet	RMSE	4.570	5.615	6.560	7.494	8.411	8.953
	MAE	3.600	4.262	4.982	5.644	6.319	6.933
	R^2	0.908	0.862	0.812	0.756	0.693	0.654
DAGCGN	RMSE	4.539	5.604	6.322	7.424	8.317	8.943
	MAE	3.577	4.353	4.887	5.604	6.288	6.774
	R^2	0.909	0.862	0.825	0.760	0.701	0.655
Causal-STAN (Ours)	RMSE	**4.180**	**5.158**	**5.972**	**7.037**	**7.963**	**8.119**
	MAE	**3.247**	**3.306**	**4.328**	**5.238**	**6.052**	**6.176**
	R^2	**0.923**	**0.916**	**0.858**	**0.798**	**0.729**	**0.708**

directionally, this contradicts the causality principle of time series prediction. Additionally, the large parameter count and complex recurrent network structure of BiLSTM may impact model efficiency and performance. ST-CausalConvNet also utilizes CNN to capture spatial features but employs a feedforward network, TCN, for processing temporal features, ensuring that causality is integrated into the temporal prediction phase. This method effectively avoids future temporal information interference, giving ST-CausalConvNet a slight edge in predictive performance. This demonstrates that in PM2.5 prediction, feedforward predictive models and consideration of causality have distinct advantages. However, ST-CausalConvNet is limited in capturing long-distance dependencies in the time series and extracting internal correlation information from input data. In comparisons of predicting PM2.5 concentrations over the next 6 h, DAGCGN's performance surpasses that of ST-CausalConvNet, thus exhibiting the best predictive performance among all baseline models. Unlike other baseline models, DAGCGN utilizes an enhanced GCN network combined with a GRU framework to effectively learn complex spatio-temporal features between sites. Yet, its limitation lies in the complex structure of the GCN network used during the spatial feature extraction phase. While this network adeptly captures spatial correlations between sites, it is restricted to processing a fixed number of related site features, allowing only a predefined number K of related sites per target site. Moreover, it is noteworthy that these baseline models do not account for the domain knowledge of air pollution dispersion or the directional characteristics of PM2.5 concentration dispersion between stations.

Compared to DAGCGN, Causal-STAN reduces the RMSE by 7.91%, the MAE by 9.23%, and improves $\boldsymbol{R^2}$ by 1.54% for the next hour's prediction. In the sixth-hour prediction, $\boldsymbol{R^2}$ for Causal-STAN remains at 0.708, still better than DAGCGN's 0.655, with its RMSE and MAE also 9.21% and 8.82% lower, respectively. This outstanding performance is attributed to the unique architectural design of Causal-STAN. For spatial feature extraction, we employ CNNs to identify temporal local correlation trends and spatial correlations within the spatio-temporal data, and we introduce directional residual blocks to extract directional characteristics of PM2.5 concentration dispersion between sites. Compared to DAGCGN's spatial feature extraction module, our module can conveniently handle model inputs with varying numbers of related sites. In terms of temporal feature processing, instead of using complex recurrent neural networks, we propose a more streamlined feedforward prediction network. By integrating TCN with an attention mechanism, we effectively capture the internal correlation information and long-term dependencies within the PM2.5 time series, providing an efficient alternative to recurrent network architectures.

To specifically showcase the predictive performance at a single target site, we selected site number 1141 and displayed its forecasting

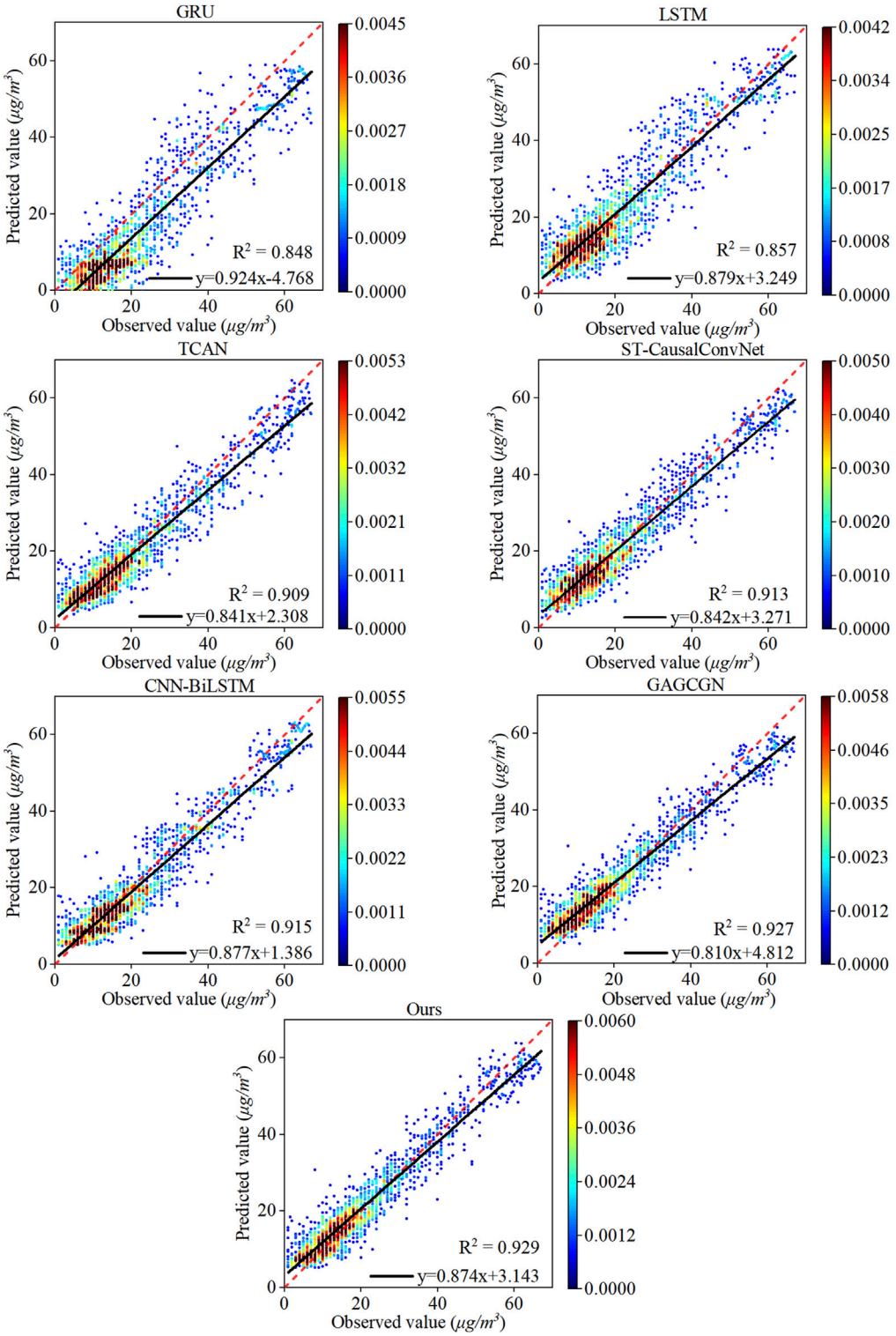


FIGURE 8
Correlation analysis between observed and predicted PM2.5 concentrations from various models on the test set at site number 1141. The dashed line is the y = x reference line and the solid line is the regression line.

results on the test dataset, which consists of 1,749 data points, in a scatter plot format. Figure 8 presents the comparison of the linear correlation between the predicted and observed PM2.5 concentrations using our model, Causal-STAN, and six other models at this site. The results demonstrate that our model attains the highest correlation coefficient (R^2 = 0.929) and a better fit between the predicted regression line and the observed reference line.

TABLE 3 Comparative performance analysis of causal-STAN and causal-STAN-year models at four monitoring sites.

Model	Site 1233			Site 1270			Site 1155			Site 1141		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
Causal-STAN (Ours)	3.189	2.224	0.951	5.034	3.964	0.929	5.063	3.783	0.922	3.985	3.088	0.929
Causal-STAN-year	3.758	2.754	0.928	5.606	4.188	0.909	6.347	4.877	0.882	5.014	3.814	0.912

3.3 Comparative evaluation of dataset partitioning strategies

To assess the effectiveness of our dataset partitioning strategy within the experimental framework, we conducted comparative experiments at four sample monitoring sites (IDs 1233, 1270, 1155, and 1141), contrasting our monthly partitioning approach (Causal-STAN) with the annual partitioning method commonly used in many models. Specifically, in the Causal-STAN method, the dataset is partitioned monthly with the first 60% designated for training, the subsequent 20% for validation, and the final 20% for testing. Conversely, in the comparative method (Causal-STAN-year), the dataset is divided annually, maintaining a consistent distribution of 60% for training, 20% for validation, and 20% for testing. This involves using data from the first 8 months of the year for training, data from September and October for validation, and the final 2 months' data for testing. Throughout the experiment, both approaches were evaluated using the same model architecture, hyperparameters, and training settings.

Table 3 presents the forecasting results obtained at various test sites for both comparative methods. The Causal-STAN model consistently showed lower RMSE and MAE values across all sites, indicating a higher accuracy in predicting PM2.5 concentrations compared to the Causal-STAN-year model. For example, at site 1233, the RMSE for Causal-STAN was 3.189, compared to 3.758 for Causal-STAN-year, representing a 15.1% decrease. Similarly, the MAE decreased from 2.754 to 2.224, a reduction of about 19.2%. Furthermore, the $\boldsymbol{R^2}$ values, which measure the proportion of variance in the dependent variable that can be predicted from the independent variables, were also higher for Causal-STAN. Specifically, at site 1233, $\boldsymbol{R^2}$ improved from 0.928 with Causal-STAN-year to 0.951 with Causal-STAN, indicating a 2.5% increase in the model's explanatory power and demonstrating a better fit to the observed data. At site 1155, the performance disparity was even more pronounced. The $\boldsymbol{R^2}$ value increased by 4.54%, underscoring the enhanced adaptability and predictive capacity of the monthly partitioning model under varied environmental conditions, which is crucial for accurate air quality forecasting.

The superior performance of the Causal-STAN model is primarily due to its ability to effectively capture short-term fluctuations and seasonal trends. The model's monthly partitioning strategy ensures comprehensive inclusion of a wide array of critical environmental variables throughout the year during the training process. This granular approach provides the essential details necessary for accurately predicting changes in PM2.5 concentrations. Moreover, by integrating data from each month, the Causal-STAN model consistently represents the unique environmental characteristics and conditions of all seasons. This capability is particularly crucial for pollutants like PM2.5, which are highly sensitive to both seasonal variations and episodic changes. Consequently, this significantly enhances the model's accuracy and robustness, enabling it to perform effectively across various times of the year. In contrast, the Causal-STAN-year model, while demonstrating reasonable performance during training and validation, exhibits a significant increase in error rates during the testing phase. This suggests an overfitting problem, where the model performs well on familiar data but struggles to adapt to new, unseen data. The annual partitioning contributes to this by smoothing over crucial short-term variations and anomalies within the dataset. Furthermore, this partitioning method may overlook critical PM2.5 variation patterns, such as lower concentrations in summer and higher in winter. Since the model is trained predominantly with summer data and tested on winter data, it fails to account for these seasonal trends. This explains why, despite effective training on large data blocks from the initial months, the model fails to accurately predict data from later months.

TABLE 4 Performance metrics of causal-STAN at various levels and kernel sizes using RMSE at monitoring site 1141, with best results highlighted in bold.

		Kernel sizes						
		2	3	4	5	6	7	8
Levels	2	5.844	4.345	4.458	4.402	4.458	4.797	4.910
	3	5.417	4.117	4.025	4.214	4.547	5.206	5.864
	4	5.483	4.125	**3.985**	4.106	4.319	4.532	4.745
	5	5.556	4.221	4.139	4.196	4.588	4.801	4.937
	6	4.945	4.397	4.322	4.717	4.636	4.824	4.914
	7	5.702	4.347	4.178	4.373	4.430	4.950	5.063
	8	5.668	4.618	4.493	4.550	4.957	5.183	5.775

3.4 Hyperparameter study

In the temporal causal attention convolutional network, we investigated the effects of two crucial hyperparameters—number of layers (levels) and kernel size—on the Causal-STAN model using a grid search strategy. Experiments were performed at monitoring site number 1141. In this context, "levels" denotes the number of layers within the network architecture, reflecting the depth of the network, whereas "kernel size" indicates the dimension of filters within the network. Table 4 presents the performance of these parameters across various combinations, utilizing RMSE as the performance metric. The lowest RMSE (3.985) occurred when both parameters were set to 4, with minimal error variance

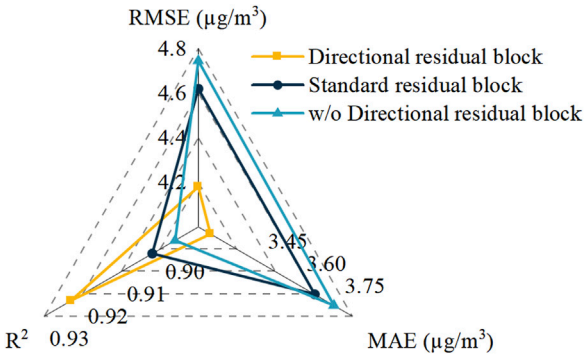


FIGURE 9
Impact of the directional residual block on prediction performance.

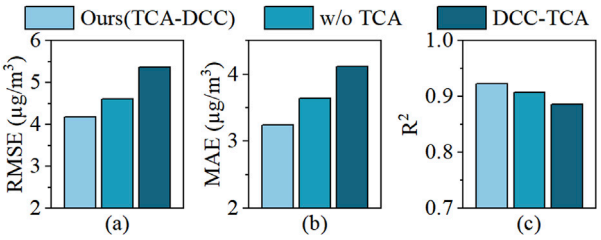


FIGURE 10
Impact of temporal causal attention on prediction performance: **(A)** RMSE, **(B)** MAE, and **(C)** R^2 comparisons for different models.

around this configuration, highlighting that balanced selections of layers and kernel size are crucial for optimal performance.

From the data in Table 4, it is evident that both overly large and small parameter values negatively affect model performance. For instance, when kernel sizes exceed 6 or levels surpass 6, RMSE increases noticeably, indicating overfitting or insufficient model generalization. Conversely, smaller values for these hyperparameters may lead to underfitting, as the network is unable to capture complex spatio-temporal relationships in the data.

The degradation in model accuracy stems from several factors. Too many layers increase the model's complexity, which complicates optimization and risks overfitting. In contrast, an insufficient number of layers fails to capture deeper data relationships, leading to underfitting. Similarly, excessively large kernel sizes may include excessive historical information, introducing noise, while too small kernel sizes may not capture sufficient long-term dependencies.

As shown in Table 4, the optimal range for both levels and kernel size lies between 3 and 5. This range provides a balance between model complexity and generalization capability, ensuring more robust performance across different data conditions. These results highlight the importance of considering the combined effects of hyperparameters to maintain model accuracy.

4 Discussion

4.1 Effects of directional residual block

To validate the effectiveness of the directional residual block, two comparison strategies were employed: a) removing the directional residual block; b) substituting the directional residual block with a standard residual block (3 × 3 convolution). The outcomes, which reflect average data across all monitoring sites in the dataset, are presented in Figure 9. Initially, removing the directional residual block led to an increase in RMSE and MAE from 4.180 and 3.247 to 4.746 and 3.728, respectively, marking increases of 13.54% and 14.82%. Concurrently, $\boldsymbol{R^2}$ declined from 0.923 to 0.896, a decrease of 2.93%. This highlights the importance of incorporating domain knowledge in multi-site PM2.5 prediction. Conversely, incorporating standard residual blocks did not markedly enhance prediction performance. This indicates that the directional residual block effectively captures the directional characteristics of PM2.5 concentration diffusion between sites, and incorporating such domain knowledge is practically beneficial for PM2.5 prediction tasks.

4.2 Effects of temporal causal attention

To evaluate the effectiveness of temporal causal attention for capturing temporal dependence, two comparison strategies were utilized: a) the removal of temporal causal attention, and b) the positioning of temporal causal attention after dilated causal convolution (DCC-TCA). The outcomes, which reflect average data across all monitoring sites in the dataset, are presented in Figure 10. Experimental results show a decrease in RMSE and MAE from 4.610 and 3.636 to 4.180 and 3.247, respectively, representing decreases of 9.33% and 10.70%, with $\boldsymbol{R^2}$ improving by 1.65% following the introduction of temporal causal attention. This underscores the importance of considering the time series' long-term dependence.

Positioning temporal causal attention after dilated causal convolution significantly diminishes the model's predictive performance. Compared to the proposed temporal causal attention convolutional network, there is a significant increase in RMSE and MAE, and a notable decrease in $\boldsymbol{R^2}$ by 4.01%. This phenomenon could be attributed to dilated causal convolution losing or weakening critical temporal information during computation, hindering temporal causal attention from fully compensating for this loss. Hence, this confirms the superiority of our proposed temporal causal attention convolutional network design.

5 Conclusion

This study introduces Causal-STAN, a novel spatio-temporal attention causal convolutional neural network, tailored for multi-site PM2.5 prediction. It addresses prevalent issues in existing methods, such as over-reliance on recurrent networks for temporal dependency extraction, limited exploration of feedforward modeling, and neglect of air quality pollutant dispersion domain knowledge. Causal-STAN integrates a multi-site spatio-temporal feature integration module and a temporal

causal attention convolutional network, significantly enhancing spatial and temporal dependency learning while introducing causality into temporal dependency analysis. Combining a convolutional neural network with a directional residual block, Causal-STAN effectively captures directional features of PM2.5 dispersion and utilizes dilated causal convolution and attention mechanisms as viable alternatives to recurrent networks for long-term dependency learning. Research results indicate that Causal-STAN can accurately forecast PM2.5 concentrations across multiple monitoring station areas for the next 6 hours, outperforming current methodologies. Specifically, its application in forecasting future PM2.5 concentrations for 247 air quality monitoring stations in the Yangtze River Delta region of China will significantly assist policymakers in more effectively evaluating and addressing air quality issues. This enhancement in predictive capability is crucial for improving public health protection and mitigating health risks associated with air pollution.

Despite the encouraging results, the Causal-STAN model has several limitations. So far, the model has only been applied to the prediction of PM2.5, and its generalizability to other pollutants, such as NO_2 or O_3, remains uncertain. Additionally, the model has only been tested in the Yangtze River Delta region, and its performance in areas with different climatic or environmental conditions has yet to be validated. The model also relies on data from existing air quality monitoring stations, which limits its effectiveness in regions with scarce monitoring infrastructure. Furthermore, data gaps are currently filled using traditional interpolation techniques, which may not fully capture the complexity of missing data. Finally, external factors such as industrial emissions or policy changes, which may occur unexpectedly, have not yet been considered, and these factors could potentially affect the model's predictive performance.

In the future, we plan to enhance our model along three targeted dimensions:

(1) Data Completion Using Deep Learning: We will integrate advanced deep learning methods, such as generative adversarial networks (GANs) or autoencoders, to perform more accurate data imputation. Preliminary experiments using a small subset of incomplete data suggest that these techniques can significantly improve the model's performance in regions with data gaps. We expect that adopting these methods will enhance overall predictive accuracy, especially in areas where air quality monitoring stations have significant data missing for certain periods.

(2) Geographical Expansion: We plan to obtain extensive pollutant data and air quality indices from the past year for the Jing-Jin-Ji or Pearl River Delta regions through the urban air quality real-time publishing platform of the China National Environmental Monitoring Center, as well as corresponding meteorological data from the National Climatic Data Center. Using data from these regions, we aim to validate the model's predictive performance under different climatic and environmental conditions. These results will guide us in refining the model for better generalization and applicability across diverse geographical areas.

(3) Application in Unmonitored Areas: We plan to collect additional data sources for unmonitored areas through open-source APIs, including functional area classification, road network data, weather forecast data, and pollutant and meteorological data from the nearest air quality monitoring stations. Furthermore, we will improve the model's training strategy with a focus on simulating the interactions between different monitoring stations. By leveraging these enriched data sources, we aim to predict pollutant levels in unmonitored areas, addressing a critical gap in air quality management and providing stronger support for public health initiatives.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YW: Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing–review and editing. ST: Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing–original draft. PZ: Conceptualization, Formal Analysis, Validation, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We thank the faculty and staff of the School of Computer Science and Technology for their support and assistance in providing the necessary resources and environment to conduct our research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ai, S., Wang, C., Qian, Z., Cui, Y., Liu, Y., Acharya, B. K., et al. (2019). Hourly associations between ambient air pollution and emergency ambulance calls in one central Chinese city: implications for hourly air quality standards. *Sci. Total Environ.* 696, 133956. doi:10.1016/j.scitotenv.2019.133956

Bi, J., Zhang, X., Yuan, H., Zhang, J., and Zhou, M. (2022). A hybrid prediction method for realistic network traffic with temporal convolutional network and LSTM. *IEEE Trans. Automation Sci. Eng.* 19, 1869–1879. doi:10.1109/TASE.2021.3077537

Chen, M.-H., Chen, Y.-C., Chou, T.-Y., and Ning, F.-S. (2023). PM2.5 concentration prediction model: a CNN–rf ensemble framework. *Int. J. Environ. Res. Public Health* 20, 4077. doi:10.3390/ijerph20054077

Chiang, P.-W., and Horng, S.-J. (2021). Hybrid time-series framework for daily-based PM $_{2.5}$ forecasting. *IEEE Access* 9, 104162–104176. doi:10.1109/ACCESS.2021.3099111

Chinatamby, P., and Jewaratnam, J. (2023). A performance comparison study on PM2.5 prediction at industrial areas using different training algorithms of feedforward-backpropagation neural network (FBNN). *Chemosphere* 317, 137788. doi:10.1016/j.chemosphere.2023.137788

de Hoogh, K., Héritier, H., Stafoggia, M., Künzli, N., and Kloog, I. (2018). Modelling daily PM2.5 concentrations at high spatio-temporal resolution across Switzerland. *Environ. Pollut.* 233, 1147–1154. doi:10.1016/j.envpol.2017.10.025

Du, S., Li, T., Yang, Y., and Horng, S.-J. (2021). Deep air quality forecasting using hybrid deep learning framework. *IEEE Trans. Knowl. Data Eng.* 33, 2412–2424. doi:10.1109/TKDE.2019.2954510

Erden, C. (2023). Genetic algorithm-based hyperparameter optimization of deep learning models for PM2.5 time-series prediction. *Int. J. Environ. Sci. Technol.* 20, 2959–2982. doi:10.1007/s13762-023-04763-6

Faraji, M., Nadi, S., Ghaffarpasand, O., Homayoni, S., and Downey, K. (2022). An integrated 3D CNN-GRU deep learning method for short-term prediction of PM2.5 concentration in urban environment. *Sci. Total Environ.* 834, 155324. doi:10.1016/j.scitotenv.2022.155324

Gao, X., Huang, J., Cardenas, A., Zhao, Y., Sun, Y., Wang, J., et al. (2022). Short-term exposure of PM $_{2.5}$ and epigenetic aging: a quasi-experimental study. *Environ. Sci. Technol.* 56, 14690–14700. doi:10.1021/acs.est.2c05534

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., et al. (2005). Fully coupled "online" chemistry within the WRF model. *Atmos. Environ.* 39, 6957–6975. doi:10.1016/j.atmosenv.2005.04.027

Hao, H., Wang, Y., Xue, S., Xia, Y., Zhao, J., and Shen, F. (2020). Temporal convolutional attention-based network for sequence modeling. *arXiv Prepr. arXiv: 2002.12530*. doi:10.48550/arXiv.2002.12530

Khandelwal, U., He, H., Qi, P., and Jurafsky, D. (2018). "Sharp nearby, fuzzy far away: how neural language models use context," in *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers)* (Stroudsburg, PA, USA: Association for Computational Linguistics), 284–294. doi:10.18653/v1/P18-1027

Lagesse, B., Wang, S., Larson, T. V., and Kim, A. A. (2020). Predicting PM $_{2.5}$ in well-mixed indoor air for a large office building using regression and artificial neural network models. *Environ. Sci. Technol.* 54, 15320–15328. doi:10.1021/acs.est.0c02549

Lai, X., Li, H., and Pan, Y. (2021). A combined model based on feature selection and support vector machine for PM2.5 prediction. *J. Intelligent and Fuzzy Syst.* 40, 10099–10113. doi:10.3233/JIFS-202812

Li, W., Wei, Y., An, D., Jiao, Y., and Wei, Q. (2022). LSTM-TCN: dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. *Environ. Sci. Pollut. Res.* 29, 39545–39556. doi:10.1007/s11356-022-18914-8

Liang, J., and Tang, W. (2022). Ultra-short-term spatiotemporal forecasting of renewable resources: an attention temporal convolutional network-based approach. *IEEE Trans. Smart Grid* 13, 3798–3812. doi:10.1109/TSG.2022.3175451

Liang, Y., Xia, Y., Ke, S., Wang, Y., Wen, Q., Zhang, J., et al. (2023). AirFormer: predicting nationwide air quality in China with transformers. *Proc. AAAI Conf. Artif. Intell.* 37, 14329–14337. doi:10.1609/aaai.v37i12.26676

Liu, B., Tan, X., Jin, Y., Yu, W., and Li, C. (2021). Application of RR-XGBoost combined model in data calibration of micro air quality detector. *Sci. Rep.* 11, 15662. doi:10.1038/s41598-021-95027-1

Liu, H., and Chen, C. (2020). Prediction of outdoor PM2.5 concentrations based on a three-stage hybrid neural network model. *Atmos. Pollut. Res.* 11, 469–481. doi:10.1016/j.apr.2019.11.019

Liu, Z., Jiang, P., De Bock, K. W., Wang, J., Zhang, L., and Niu, X. (2024). Extreme gradient boosting trees with efficient Bayesian optimization for profit-driven customer churn prediction. *Technol. Forecast Soc. Change* 198, 122945. doi:10.1016/j.techfore.2023.122945

Marsha, A., and Larkin, N. K. (2019). A statistical model for predicting PM $_{2.5}$ for the western United States. *J. Air Waste Manage Assoc.* 69, 1215–1229. doi:10.1080/10962247.2019.1640808

Marvin, D., Nespoli, L., Strepparava, D., and Medici, V. (2022). A data-driven approach to forecasting ground-level ozone concentration. *Int. J. Forecast* 38, 970–987. doi:10.1016/j.ijforecast.2021.07.008

Nasr Azadani, M., and Boukerche, A. (2022). A novel multimodal vehicle path prediction method based on temporal convolutional networks. *IEEE Trans. Intelligent Transp. Syst.* 23, 25384–25395. doi:10.1109/TITS.2022.3151263

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., et al. (2011). Detecting novel associations in large data sets. *Science* 334, 1518–1524. doi:10.1126/science.1205438

Shakya, D., Deshpande, V., Goyal, M. K., and Agarwal, M. (2023). PM2.5 air pollution prediction through deep learning using meteorological, vehicular, and emission data: a case study of New Delhi, India. *J. Clean. Prod.* 427, 139278. doi:10.1016/j.jclepro.2023.139278

Sun, Y., Ding, J., Liu, Z., and Wang, J. (2023). Combined forecasting tool for renewable energy management in sustainable supply chains. *Comput. Ind. Eng.* 179, 109237. doi:10.1016/j.cie.2023.109237

Tariq, S., Tariq, S., Kim, S., Woo, S. S., and Yoo, C. (2023). Distance adaptive graph convolutional gated network-based smart air quality monitoring and health risk prediction in sensor-devoid urban areas. *Sustain Cities Soc.* 91, 104445. doi:10.1016/j.scs.2023.104445

Teng, M., Li, S., Yang, J., Wang, S., Fan, C., Ding, Y., et al. (2023). Long-term PM2.5 concentration prediction based on improved empirical mode decomposition and deep neural network combined with noise reduction auto-encoder- A case study in Beijing. *J. Clean. Prod.* 428, 139449. doi:10.1016/j.jclepro.2023.139449

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *arXiv Prepr. arXiv:1706.03762*. doi:10.48550/arXiv.1706.03762

Wang, H., Zhang, L., Wu, R., and Cen, Y. (2023). Spatio-temporal fusion of meteorological factors for multi-site PM2.5 prediction: a deep learning and time-variant graph approach. *Environ. Res.* 239, 117286. doi:10.1016/j.envres.2023.117286

Wang, J., Wang, D., Zhang, F., Yoo, C., and Liu, H. (2024a). Soft sensor for predicting indoor PM2.5 concentration in subway with adaptive boosting deep learning model. *J. Hazard Mater* 465, 133074. doi:10.1016/j.jhazmat.2023.133074

Wang, J., Wu, T., Mao, J., and Chen, H. (2024b). A forecasting framework on fusion of spatiotemporal features for multi-station PM2.5. *Expert Syst. Appl.* 238, 121951. doi:10.1016/j.eswa.2023.121951

Wang, P., Zhang, H., Qin, Z., and Zhang, G. (2017). A novel hybrid-Garch model based on ARIMA and SVM for PM 2.5 concentrations forecasting. *Atmos. Pollut. Res.* 8, 850–860. doi:10.1016/j.apr.2017.01.003

Wang, S., Li, Y., Zhang, J., Meng, Q., Meng, L., and Gao, F. (2020). "PM2.5-GNN," in *Proceedings of the 28th international conference on advances in geographic information systems* (New York, NY, USA: ACM), 163–166. doi:10.1145/3397536.3422208

Wu, E. M.-Y., and Kuo, S.-L. (2012). Air quality time series based GARCH model analyses of air quality information for a total quantity control district. *Aerosol Air Qual. Res.* 12, 331–343. doi:10.4209/aaqr.2012.03.0051

Yan, R.-H., Peng, X., Lin, W., He, L.-Y., Wei, F.-H., Tang, M.-X., et al. (2022). Trends and challenges regarding the source-specific health risk of PM $_{2.5}$ -bound metals in a Chinese megacity from 2014 to 2020. *Environ. Sci. Technol.* 56, 6996–7005. doi:10.1021/acs.est.1c06948

Yang, Z., Mahendran, R., Yu, P., Xu, R., Yu, W., Godellawattage, S., et al. (2022). Health effects of long-term exposure to ambient PM2.5 in asia-pacific: a systematic review of cohort studies. *Curr. Environ. Health Rep.* 9, 130–151. doi:10.1007/s40572-022-00344-w

Young, T., Hazarika, D., Poria, S., and Cambria, E. (2018). Recent trends in deep learning based natural language processing [review article]. *IEEE Comput. Intell. Mag.* 13, 55–75. doi:10.1109/MCI.2018.2840738

Zhang, B., Zou, G., Qin, D., Ni, Q., Mao, H., and Li, M. (2022). RCL-Learning: ResNet and convolutional long short-term memory-based spatiotemporal air pollutant concentration prediction model. *Expert Syst. Appl.* 207, 118017. doi:10.1016/j.eswa.2022.118017

Zhang, L., Na, J., Zhu, J., Shi, Z., Zou, C., and Yang, L. (2021). Spatiotemporal causal convolutional network for forecasting hourly PM2.5 concentrations in Beijing, China. *Comput. Geosci.* 155, 104869. doi:10.1016/j.cageo.2021.104869

Zhou, H., Zhang, F., Du, Z., and Liu, R. (2021). Forecasting PM2.5 using hybrid graph convolution-based model considering dynamic wind-field to offer the benefit of spatial interpretability. *Environ. Pollut.* 273, 116473. doi:10.1016/j.envpol.2021.116473

Zhu, J., Deng, F., Zhao, J., and Zheng, H. (2021). Attention-based parallel networks (APNet) for PM2.5 spatiotemporal prediction. *Sci. Total Environ.* 769, 145082. doi:10.1016/j.scitotenv.2021.145082

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 11 November 2024
DOI 10.3389/fenvs.2024.1457801

OPEN ACCESS

EDITED BY
Sayali Apte,
Symbiosis International University, India

REVIEWED BY
Carmen Cantuarias-Villessuzanne,
ESPI2R Research in Real Estate, France
Gautam Siddharth Kashyap,
Jamia Hamdard University, India

*CORRESPONDENCE
Xiang Ren,
rx17311986636@qlit.edu.cn

RECEIVED 01 July 2024
ACCEPTED 29 October 2024
PUBLISHED 11 November 2024

CITATION
Peng W, Li J and Ren X (2024) Can smart city construction be the answer to sustainable development? Evidence from Chinese corporates.
Front. Environ. Sci. 12:1457801.
doi: 10.3389/fenvs.2024.1457801

Can smart city construction be the answer to sustainable development? Evidence from Chinese corporates

Weihui Peng[1], Jie Li[1] and Xiang Ren[2]*

[1]School of Economics, Guizhou University of Finance and Economics, Guiyang, China, [2]Business School, Qilu Institute of Technology, Jinan, China

The construction of smart cities plays a pivotal role in promoting regional sustainability by utilizing technology-driven urban development in the digital era. This study employs the difference-in-differences model to empirically analyze the impact of the smart city pilot policy on corporate carbon intensity, using data from Chinese A-share listed corporates from 2009 to 2021. The findings are as follows: First, the smart city pilot policy significantly reduces corporate carbon intensity in pilot cities, and this conclusion remains robust after a series of sensitivity tests. Second, the policy exhibits heterogeneous effects on corporate carbon intensity across different industries and city locations, with more pronounced effects observed in central cities, traditional industries, and heavily polluting industries. Third, mechanism analysis reveals that the policy reduces corporate carbon intensity through three channels: promoting technological innovation, increasing external market attention, and providing ex-ante government subsidies.

KEYWORDS

smart city, sustainability, corporates carbon intensity, difference-in-differences, mechanism

1 Introduction

Climate change presents an urgent and persistent threat to humanity. The continued rise of greenhouse gas emissions is generating widespread and significant adverse effects on global agricultural yields, socioeconomic stability, and human well-being, ultimately jeopardizing the sustainable development of human society. At present, the Chinese government attaches great importance to addressing climate change and is actively implementing measures to conserve energy and reduce emissions. Enterprises are not only the main body of market economic activities, but also the main implementors of energy saving and carbon emission reduction (Haney, 2017; Chen et al., 2022). As a key factor for enterprises to achieve carbon emission reduction breakthrough, green technology innovation is significantly affected by urban strategy (Xie et al., 2021; Feng et al., 2024). Smart city pilots promote sustainable urban development by building digital infrastructure and knowledge ecosystems (Shen et al., 2023). The policy promotes green technology innovation in enterprises by enhancing the level of digitization in enterprises. In 2008, IBM introduced the concept of a "Smarter Planet," which significantly influenced China's approach to urban development. On 5 December 2012, the General Office of China's Ministry of Housing and Urban-Rural Development formally issued the "Notice on Launching Pilot Projects for National Smart Cities", along with the "Interim

Management Measures for the Pilot Projects of National Smart Cities" and the "Indicator System for the Pilot Projects of National Smart Cities (Trial)". Between 2012 and 2015, China launched three batches of pilot smart cities involving over 300 counties and cities. The pilot program clearly required that cities applying for pilot status should use the indicator system as a reference to develop achievable goals and implementation plans tailored to their local context. The indicator system encompasses various aspects such as security systems, network infrastructure, public platforms and databases, intelligent construction and livability, smart management and services, as well as smart industries and economy. In 2014, the State Council issued the "National New Urbanization Plan (2014–2020)", elevating smart city development to a strategic level and outlining six key directions: broadband information networks, informatized planning and management, intelligent infrastructure, convenient public services, modernized industrial development, and refined social governance. This research examines the intersection of smart city initiatives and corporate sustainability: Do smart city pilot programs, implemented alongside global energy and emissions reduction goals, tangibly influence corporate carbon reduction endeavors? What policy levers drive this relationship?

Mitigating carbon dioxide emissions presents a pressing challenge for both contemporary society and the corporate sector. The escalating concentration of carbon dioxide, a key driver of global warming, presents a multitude of challenges to human development. These include economic stagnation, adverse health impacts, a heightened frequency of extreme weather events, and rising sea levels (Mora et al., 2018; Dahlmann et al., 2019; Liu L. et al., 2022). The topic of corporate carbon reduction has attracted widespread attention from scholars, with numerous studies exploring its various aspects. A substantial body of research examines the influence of firm-level characteristics on corporate carbon performance, with studies highlighting the roles of firm size, stakeholder pressure, and export behavior, among other factors (Sadler, 2016; Kumarasiri, 2017; Cui et al., 2016; Richter and Schiersch, 2017). Other studies examine external factors, finding that internationalization level, financial support, pollution prevention actions and institutional regulations all influence corporate participation in carbon reduction (Sadler, 2016; Wang et al., 2019; Niu et al., 2023; Feng et al., 2023).

Researchers have employed various methodologies, including double difference model, panel quantile regression model, and synthetic control method, to validate the inhibitory effect of smart city construction on carbon emissions in pilot cities (Zawieska and Pieriegud, 2018; Wang et al., 2021; Guo et al., 2022). In terms of mechanisms, smart cities foster digital technology innovation that directly contributes to reducing carbon emissions (Jiang et al., 2021), while also facilitating urban carbon reduction through industrial upgrading, enhancing energy efficiency, bolstering carbon absorption capacity among other measures (Caragliu and Del Bo, 2019; Chu et al., 2021; Shen et al., 2024). Smart city construction not only promotes inclusive financial development and addresses energy poverty issues within enterprises to curtail enterprise-level carbon emissions (Gao and Yuan, 2022), but also reduces household transportation and education-related emissions through the development of smart technology (Wu, 2022). Some scholars categorize smart city pilot policies into three types: smart governance, smart industry, and smart livelihood, suggesting that smart industry policies reduce carbon emissions by driving data aggregation and optimizing industrial structures, while smart governance and smart livelihood policies promote carbon reduction by stimulating green technology innovation and enhancing energy efficiency (Nicolas et al., 2021). It is evident that existing research still requires further exploration in several areas. First, the existing literature lacks a comprehensive framework that integrates both smart city development and corporate carbon emissions intensity. There is a notable absence of research exploring the theoretical underpinnings and inherent mechanisms through which smart city initiatives influence a firm's carbon footprint. Second, current studies predominantly focus on the construction of smart cities themselves, with insufficient theoretical research on how smart city construction affects corporate carbon emissions. Third, there is a dearth of empirical research that delves into the relationship between smart city development and corporate carbon emission intensity using granular, micro-level data.

The contributions of this study are threefold: Firstly, while previous studies have often relied on aggregate or *per capita* carbon emission data, this research adopts carbon emission intensity as a more precise indicator of corporate-level performance, especially given the unique dynamics of developing economies. Secondly, starting from the smart city pilot policy, a rigorous quasi-natural experimental design is employed to identify its net policy effect on corporate carbon emission intensity. This clarifies the mechanism by which the smart city pilot policy helps reduce corporate carbon emission intensity, revealing specific pathways through green technology innovation, external attention, and government subsidies. Specifically, smart city development promotes the digital economy, improves the efficiency of resource allocation such as talent, capital, and land, increases R&D investment, and fosters corporate green technology innovation (Yang et al., 2024; Guo et al., 2024). It also enhances investors' environmental awareness (Guo et al., 2024), improves corporate information transparency (Mahmood et al., 2020), and strengthens market external supervision capabilities (Liu et al., 2020), thereby increasing market external attention to enterprises. Carbon emission is an important aspect of smart city development (Li, 2022), and the improvement of urban digitalization and informatization levels also enhances the efficiency of government subsidies, thereby reducing corporate carbon emission intensity. Thirdly, to ensure robust findings, this study carefully accounts for potential confounding effects arising from other policy interventions.

The remainder of this study is organized as follows: Section 2 establishes the theoretical framework and research hypotheses. Section 3 outlines the data sources and research design employed. Section 4 presents the empirical analysis. Section 5 explores potential avenues for further research. Finally, Section 6 concludes the study.

2 Theoretical analysis

2.1 Smart city construction and corporate carbon intensity

Smart cities represent an advanced stage of urban digitalization, where pilot policies focus primarily on leveraging information

technology transformations to elevate urban governance models, fostering innovation in resource allocation, emerging industries, and related technologies and products (Xin and Qu, 2019). Essentially, this involves providing support and assurance for environmental governance and emission control through technological innovation. The construction of smart cities enables local governments to enforce environmental laws using modern information technology methods. Given the emphasis on environmental sustainability in smart cities, the environmental requirements imposed by the government are often stricter than those in other ordinary cities. For instance, China's "National Smart City Pilot Indicator System" explicitly mandates that smart cities must achieve intelligent management of urban ecological environments, build city-level carbon emission monitoring platforms based on smart city infrastructure, and collect carbon emission and energy usage data accurately and in real-time. By interconnecting data across various stages and employing visualization techniques, real-time carbon emission monitoring is enabled, thereby enhancing governmental decision-making capabilities and strengthening institutional-level environmental supervision (Xu et al., 2023). Smart city pilot policies have driven enterprises towards a green and low-carbon transformation. The development of smart cities enhances digital infrastructure, allowing enterprises to access and share big data resources more conveniently, optimize production processes and resource allocation, and reduce energy consumption and waste emissions. Smart city policies also promote the formation of smart industrial clusters, providing companies with more opportunities for collaboration and innovation, facilitating industrial structure optimization and upgrading, and lowering corporate carbon emission intensity. Additionally, in policy implementation, there is a focus on developing new models of digital industrialization and industrial digitalization, encouraging enterprises to continuously explore these innovative models to enhance production and resource efficiency, thereby aiding in reducing corporate carbon emissions.

Hypothesis 1. Smart city development contributes to a decline in corporate carbon intensity.

2.2 The role of green technological innovation

The pilot policies for smart cities have provided enterprises with the motivation and opportunities for innovation by promoting green technology innovation, effectively reducing corporate carbon emissions. These policies aim to integrate advanced information technology and intelligent systems to enhance urban digitalization, improve the allocation efficiency of labor, land, and capital, and drive green technology innovation (Yang et al., 2024). In the process of building smart cities, government environmental subsidies, corporate environmental awareness, and R&D investment have facilitated green technology innovation in enterprises (Guo et al., 2024). For instance, the "smart city" pilot index system explicitly includes elements such as digital management, smart environmental protection, and innovation investment. The level of digitalization (Han et al., 2021), R&D investment (Wang et al., 2016), and environmental expenditure (Chen et al., 2020) are crucial factors influencing green technology innovation in enterprises. The 14th Five-Year Plan mentions that "smart city" pilots need to advance the construction of new technology infrastructure, the transformation and upgrading of traditional industries, and the establishment of a new generation of information infrastructure systems. This trend has accelerated the iteration of information technology, providing a favorable external environment for enterprise technology R&D (Zhan and Li, 2022). In the Pidu District of Chengdu, Sichuan Province, China, the implementation of the "149" smart construction project has comprehensively promoted the digital transformation of Chengdu enterprises, facilitating the integration of new-generation information technology with traditional enterprises. Green technology innovation can boost clean production in enterprises, enhance energy-saving efforts, reduce resource consumption from the production side, and foster new energy consumption methods. It empowers the optimization and upgrading of industrial structures, aiding the transition to low-carbon green industries, and reducing corporate production carbon emissions, thereby achieving source control of carbon emissions (Liu et al., 2020). The use of green technology innovation in the energy sector can accelerate the development of photovoltaics, wind power, and renewable energy, effectively promoting the development of the new energy sector. It facilitates the transition of the energy consumption structure towards green, low-carbon, and clean energy, directly reducing carbon dioxide emissions. Green technology innovation can effectively control decarbonization costs and provides technical support for the research and large-scale application of carbon dioxide utilization, capture, and storage technologies, creating a "technology dividend" effect that enhances carbon emission performance. In summary, this paper proposes research Hypothesis 2.

Hypothesis 2. By encouraging the adoption of green technologies, smart city initiatives contribute to lowering corporate carbon intensity.

2.3 The role of external market attention

Smart city pilot policies provide significant external impetus for enterprises to reduce carbon intensity by guiding public attention, particularly the scrutiny of investors and media. Market investors are not only concerned with short-term economic returns but are also beginning to incorporate environmental responsibility into their investment considerations. This green investment preference encourages enterprises to pay more attention to their own green development and reduce carbon emissions. In addition, as important supervisors in the capital market, media coverage creates a "watchdog effect", increasing the pressure on enterprises to expose their environmentally polluting behaviors. This effect compels non-compliant enterprises to enhance their environmental awareness and reduce carbon emissions (Ahern and Sosyura, 2015). Moreover, intelligent interconnected platforms utilizing digital technologies enhance enterprise information transparency. This enables external market players to conveniently access enterprise information and improves

supervision effectiveness. Digital technologies represented by digital supply chain platforms can extend outwards, bringing efficient communication models to enterprises and increasing the level of integration with external stakeholders (Ivanov et al., 2022). For example, Liu et al. (2023) utilized the textual content of government environmental attention to empirically examine the impact of government oversight on corporate emissions, finding that government environmental attention significantly reduced corporate carbon emissions.

Hypothesis 3. By attracting the interest of environmentally conscious investors and consumers, smart city construction encourages enterprises to embrace sustainable practices and reduce their carbon intensity.

2.4 The role of government subsidies

The smart city pilot policy effectively reduces corporate carbon emission intensity by increasing government subsidies, thereby stimulating companies' intrinsic motivation and environmental awareness. Firstly, the smart city policy considers urban carbon emissions as a crucial standard for performance assessment and governance capability evaluation (Li, 2022). As the leader in environmental governance, the government transfers part of the economic benefits to micro-market entities through financial subsidies, effectively stimulating companies' intrinsic motivation to reduce pollution and emissions (Peng and Liu, 2018). Ex-ante subsidies can be used for environmental initiatives such as upgrading environmental technologies and purchasing energy-saving and emission-reducing equipment, thereby lowering the economic costs for companies in implementing environmental measures and enhancing their enthusiasm and initiative for emission reduction. Secondly, smart city development enhances the transparency of corporate environmental information, enabling the government to more accurately monitor and assess corporate environmental performance and provide targeted ex-ante subsidies to companies with good environmental performance. Through real-time monitoring and data analysis enabled by smart city technologies, the government can more precisely identify and reward companies actively taking emission reduction measures, further incentivizing them to lower carbon emissions.

Hypothesis 4. Government subsidies provided as part of smart city construction initiatives help reduce corporate carbon intensity.

3 Methodology

3.1 Model

This paper investigates whether and to what extent smart city development influences the carbon emission intensity of enterprises. To address this question, we employ a difference-in-differences (DID) methodology, exploiting the phased roll out of the smart city pilot program. Building upon the work of Yao et al. (2020), we construct the multi-period DID model (Equation 1).

$$Intensity_{ijt} = \beta_0 + \beta_1 DID_{it} + \phi Controls_{ijt} + City_i + Firm_j + Year_t + \varepsilon_{ijt} \quad (1)$$

where the subscripts i, j and t represent the city, enterprise and time, respectively. The dependent variable *Intensity* denotes the carbon emission intensity of enterprise. The explanatory variable, denoted by *DID*, measures the implementation of the smart city pilot policy. *Controls* represents a series of control variables, whose selection will be explained in detail below. ε_{ijt} is the random error term. β_0 is the constant term and β_1 is the coefficient of the core explanatory variable, which indicates the specific impact of the smart city pilot policy on the carbon emission intensity of enterprises in the pilot areas.

3.2 Variables

3.2.1 The explained variable

The dependent variable analyzed in this study is corporate carbon emission intensity (Intensity). To construct this metric, we first collect data on firm's reported annual carbon emissions (direct, indirect, and total) from publicly accessible sources such as annual reports, social responsibility reports, and environmental reports. For firms without explicit emissions disclosures, we estimate their emissions by multiplying their reported coal consumption by corresponding industry-specific carbon emission coefficients. In reference to Chapple et al. (2013), carbon emission intensity is operationalized as carbon dioxide emissions scaled by the firm's primary business revenue.

3.2.2 The core explanatory

Our primary variable of interest is the DID estimator, operationalized as the interaction term between the treatment group dummy (Treat) and the post-policy time dummy (Post). Treat equals 1 for enterprises in cities designated as smart city pilot zones and 0 otherwise. Post takes a value of 1 for the year of policy implementation and subsequent years within the respective city, and 0 for pre-policy years.

3.2.3 The control variables

Referring to Yu et al. (2015), Wang et al. (2017), and Chen et al. (2023), this paper controls for firm-level and region-level variables that influence corporate carbon emission intensity, including firm size (Size) measured by the logarithm of total assets, leverage ratio (Lev) measured as the ratio of total liabilities to total assets, return on assets (ROA) measured as the ratio of net profit to total assets, board size (Board) measured by the logarithm of the number of board members, book-to-market ratio (BM) measured as the ratio of book value to market value, firm age (Age) represented by years since going public, growth rate of operating income (Growth) calculated as the increase in current operating income relative to previous operating income, regional economic development level (LnperGDP) indicated by logarithm *per capita* GDP, and regional financial development level (RFDL) expressed as loans balance from financial institutions relative to GDP.

3.2.4 The mediating variables

The level of green technology innovation (Innovation) is measured by the logarithm of the number of green patents obtained by enterprises. This paper examines the role of external

TABLE 1 Descriptive statistics.

Categories	Variable name	Variable code	Obs	Min	Max	Mean	Std. Dev
Explained Variable	Carbon emission intensity	Intensity	17,784	1.389	6.256	1.000	0.000
Core Explanatory Variable	Smart city pilot programmes	DID	17,784	0.000	1.000	0.466	0.318
Control Variables	Firm size	Size	17,784	20.022	26.337	1.331	22.356
	Debt-to-equity ratio	Lev	17,784	0.034	0.866	0.203	0.426
	Net profit margin on total assets	ROA	17,784	0.168	0.221	0.058	0.043
	Revenue growth rate	Growth	17,784	0.573	2.043	0.359	0.161
	Board size	Board	17,784	11.168	17.193	14.028	1.008
	Age of business	Age	17,784	1.386	3.466	0.400	2.808
	Book-to-market ratio	BM	17,784	0.141	1.180	0.247	0.642
	Regional financial development level	RFDL	17,784	0.000	3.449	0.721	1.445
	Level of economic development	LnperGDP	17,784	0.000	12.223	2.417	10.742
Mediating Variables	Green technology innovation	Innovation	17,784	0.000	3.832	0.447	0.703
	Media attention	Media-attention	17,784	0.000	30.52	2.861	4.429
	Investor attention	Investor-attention	17,784	0.000	8.922	5.965	2.436
	Government subsidies	ex-ante Subsidy	17,784	0.000	20.858	1.968	2.796
		ex-post Subsidy	17,784	0.000	23.301	16.196	7.321

market attention from two aspects: investors and media. Specifically, investor attention (Investor-attention) is measured by the logarithm of the median value of the Baidu search index for the year, which is mainly derived from the search volume of listed companies in Baidu search keywords; media attention (Media-attention) is measured by the total number of times the company was reported in online media news throughout the year. The research by Peng and Liu (2018) classifies government subsidies into two types: ex-ante and ex-post subsidies. Ex-ante government subsidies refer to subsidies that the government provides to enterprises without compensation, which is a form of ex-ante incentive. Ex-post government subsidies refer to the subsidy policy in which the government returns a certain amount of money to enterprises through tax reduction after the project reaches the expected effect or is completed, which is a form of ex-post incentive. This paper measures government ex-ante subsidies using various direct subsidies to enterprises and measures government ex-post subsidies using post-tax returns to enterprises.

3.3 Data sources

This study draws upon data from A-share listed companies on both the Shanghai and Shenzhen Stock Exchanges covering the period from 2009 to 2021. Firm-level data, including carbon emissions sourced from annual reports, environmental disclosures, and social responsibility reports, are supplemented by information from the CSMAR and Wind databases. City-level data is sourced from the China City Statistical Yearbook and the EPS database. The following data pre-processing steps were undertaken: (1) removal of companies marked as ST, ST* or PT; (2) excluding of samples from the financial sector, including banking and insurance; (3) elimination of samples with less than 3 years of data or missing values for key variables; (4) and winsorization of all continuous variables at the 1% level at both tails. The final sample comprises 1,368 listed companies, yielding a total of 17,784 observations. Among them, the manufacturing industry accounts for 14,488 observations, which is 81.47%; the mining industry accounts for 606 observations, which is 3.4%; the electricity, heat, gas, and water production and supply industries account for 881 observations, which is 4.95%; the transportation, storage, and postal services industries account for 864 observations, which is 4.86%; and other industries account for 945 observations, which is 5.31%. The descriptive statistical analysis of the data for the aforementioned variables is presented in Table 1.

4 Results

4.1 Benchmark results

We use Stata 17 as the econometric analysis tool in this paper. Table 2 presents the results of the difference-in-differences (DID) estimation. Column 1 provides a baseline estimate of the smart city pilot policy's impact on firm carbon intensity without controlling for other factors. Column 2 introduces firm, year, and city fixed effects to address potential unobserved heterogeneity across these dimensions. Column 3 builds upon this specification by incorporating both company-level and city-level control variables that may influence firm carbon intensity. The regression results consistently demonstrate a statistically significant negative relationship between the smart city pilot policy and firm carbon

TABLE 2 Benchmark results.

Variables	(1) Intensity	(2) Intensity	(3) Intensity
DID	−0.101***	−0.157***	−0.159***
	(0.016)	(0.058)	(0.057)
Size			−0.195***
			(0.033)
Lev			0.200
			(0.132)
ROA			0.370
			(0.318)
Board			−0.148
			(0.100)
Growth			0.368***
			(0.043)
Age			−0.015
			(0.155)
BM			0.682***
			(0.101)
RFDL			−0.022
			(0.052)
LnperGDP			−0.060
			(0.053)
Constant			4.847***
			(1.006)
Firm Fixed Effect	NO	YES	YES
Year Fixed Effect	NO	YES	YES
City Fixed Effect	NO	YES	YES
N	17,784	17,784	17,784
R^2	0.002	0.233	0.257

Note: (1) The figures in parentheses are clustered robust standard errors at the enterprise level; (2) *, ** and *** indicate significance at the level of 10%, 5% and 1%, respectively; (3) The same as below.

intensity. This finding holds across all three model specifications, regardless of the inclusion of control variables. This robust result suggests that the implementation of the smart city pilot policy has contributed to a significant reduction in carbon emission intensity among firms located within the pilot cities. From an economic standpoint, using the comprehensive set of control variables in Column 3, we observe an average reduction of 15.9% in the carbon intensity of firms due to the implementation of the policy. Simultaneously, the econometric results reveal that firm size (Size), revenue growth rate (Growth), and book-to-market ratio (BM) all have significant impacts on reducing corporate carbon emission intensity. Specifically, firm size (Size) significantly reduces carbon emission intensity, possibly because larger firms face stricter environmental regulations, higher requirements for environmental information disclosure, and have greater capacity to adopt green technologies for emission reduction, ultimately promoting a decrease in carbon emission intensity. Conversely, a higher revenue growth rate (Growth) is associated with higher carbon emission intensity, indicating that Chinese companies have not yet completed their green transition, which aligns with China's current stage of reaching peak carbon emissions. The book-to-market ratio (BM) has a positive impact on carbon emission intensity, suggesting that the higher the market value of a listed company (i.e., the lower the BM ratio), the more the company values green and low-carbon development and is more capable of adopting green and low-carbon technologies to transform its production model and reduce carbon emission levels.

4.2 Parallel trend test

Following Imbens and Wooldridge (2009), we construct model (Equation 2).

$$Intensity_{ijt} = \partial_0 + \sum_{t=-3}^{t=3} \partial_1 reform_{it} + \varphi Controls_{ijt} + City_i + Firm_j + Year_t + \varepsilon_{ijt} \quad (2)$$

where $Intensity_{ijt}$ is the explained variable in this paper, the variable $t < 0$ denotes the number of years preceding the implementation of the smart city pilot policy in city i, where enterprise j is located. During this period, the $reform_{it}$ variable takes on a value of 0. Similarly, when $t = 0$, it signifies the year of policy implementation. Subsequently, for $t > 0$, it indicates the number of years following the introduction of the smart city pilot policy, with the $reform_{it}$ variable assigned a value of 1. Given the extended sample period, this study focuses on illustrating the trend of the policy's effect within the initial 2–3 years following its implementation. To mitigate concerns regarding perfect multicollinearity, the third year prior to policy implementation serves as the base year, maintaining consistency throughout the remaining model specifications.

Prior to the smart city pilot policy, there was no significant difference in carbon emission intensity between enterprises in pilot and non-pilot cities (Figure 1). This confirms the necessary assumption for a valid difference-in-differences analysis. Notably, the coefficients for the year of policy implementation and subsequent years are significantly negative, indicating that the policy has had a noticeable impact on carbon emission intensity within pilot areas. This suggests that developing smart cities can contribute to long-term and sustainable reductions in local business carbon emissions.

4.3 Placebo test

To further validate our baseline regression analysis, we conduct a placebo test using a non-parametric permutation approach (Cai et al., 2016). This involves randomly assigning cities to a simulated treatment group, effectively mimicking the implementation of the smart city pilot policy in a randomly selected subset of cities instead of the designated pilot cities. By evaluating this counterfactual

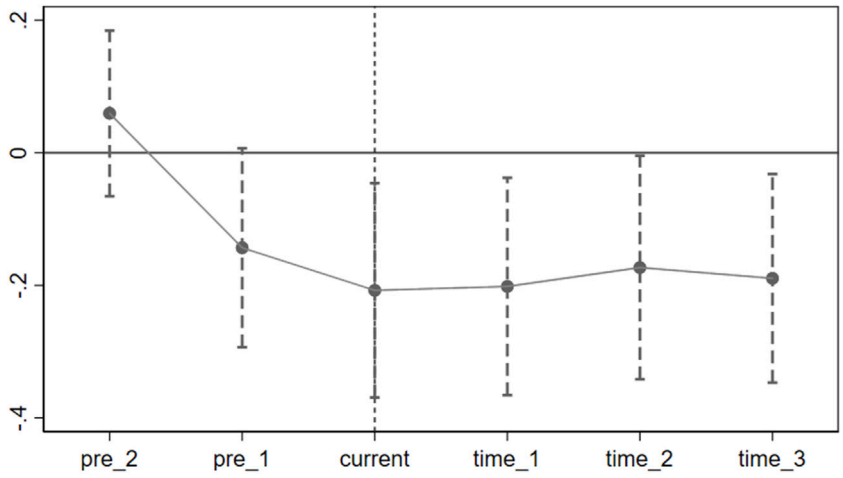


FIGURE 1
Parallel trend test.

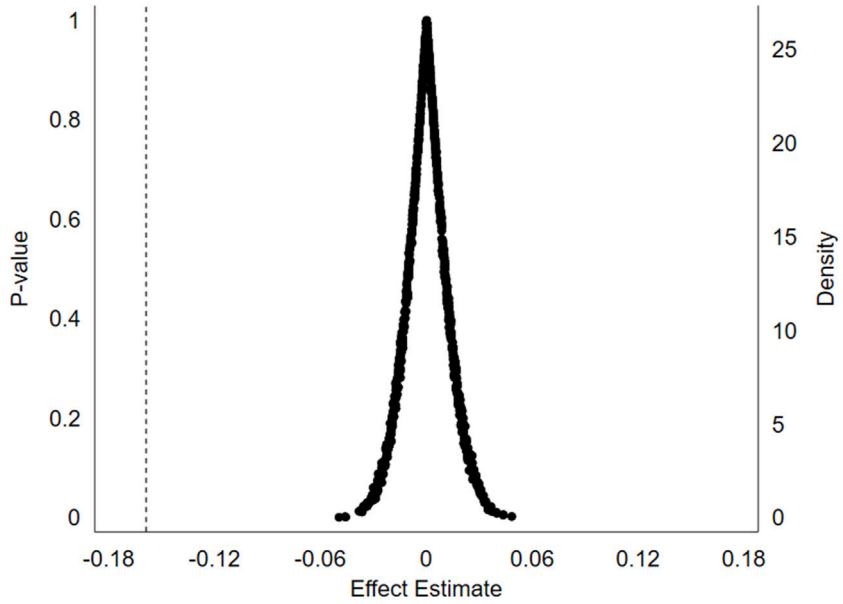


FIGURE 2
Placebo test.

scenario, we can assess the sensitivity of our initial findings to the specific selection of pilot cities and strengthen the causal relationship between the policy and observed reductions in carbon emission intensity.

Figure 2 presents the results of the placebo test, which encompassed 1,000 Monte Carlo simulations. The vertical dashed line in the figure represents the estimated DID coefficients from our baseline regression. As evident from the figure, the estimated coefficient probability density within the placebo test exhibits an approximately normal distribution. Importantly, these estimated coefficients are consistently situated far from the baseline regression results. Furthermore, a substantial majority of the *p*-values obtained from the placebo test surpass the *p*-values associated with the DID estimated coefficients in our baseline regression.

Therefore, we can confidently conclude that the baseline regression results of our study have successfully passed the placebo test. This outcome strongly supports the reliability and robustness of our findings, effectively mitigating concerns regarding spurious correlations or random fluctuations as potential drivers of our observed results. The placebo test serves as compelling evidence that our findings are likely to be genuine and not artifacts of our research design.

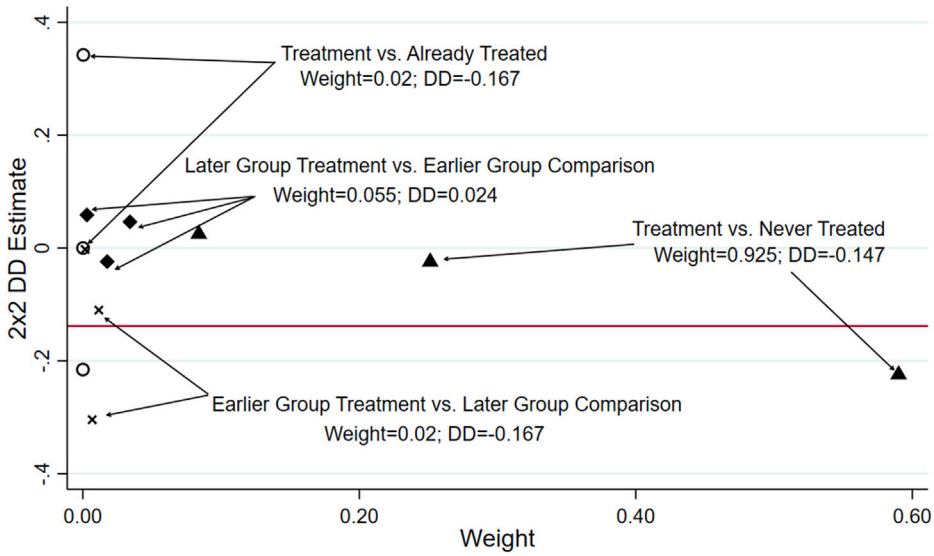


FIGURE 3
Bacon decomposition results.

4.4 Robustness checks

4.4.1 Bacon decomposition

For staggered DID models, Goodman-Bacon (2021) points out that earlier treatment groups are not good control groups (also called "bad control groups") because their pre-trends changed compared with later-treatment groups or never-treated groups (also called "good control groups"). The traditional two-way fixed effects model can produce potential bias when conducting staggered DID estimations precisely because of the existence of these "bad control groups". Therefore, this paper adopts the Bacon decomposition method to decompose the estimator into weighted values of each part to test this issue.

The decomposition of the estimated treatment effect is visually represented in Figure 3. The figure reveals that the policy has a significant net effect on reducing carbon intensity in pilot cities, accounting for 94.5% of the total estimated effect. In contrast, only 5.5% can be attributed to "bad control groups". This stark contrast highlights the high credibility of our baseline regression results. In our analysis, we utilize a "never-treated" group as the control for the period under examination. This makes our estimates of the policy's effects more reliable and helps us accurately measure the overall impact of the policy over time.

4.4.2 PSM-DID

To address potential sample selection bias, we employed a Propensity Score Matching-Difference-in-Differences (PSM-DID) approach, utilizing three matching techniques (1:2 nearest-neighbor, caliper, and kernel density) to refine the control group. This ensured comparability between treatment and control groups.

Subsequent PSM-DID estimation confirmed a statistically significant and negative impact of smart city pilot policies on corporate carbon intensity, consistent with our baseline regression findings (Table 3). This result suggests minimal sample selection bias, further strengthening the robustness of our initial conclusions.

4.4.3 Controlling competitive policy interference

During the sample selection period, the implementation of other related policies could lead to coefficient bias in the core explanatory variables. Many policies across different regions occurred simultaneously or overlapped with the smart city pilot policy, causing a certain degree of policy overlap effect and potentially undermining the validity of regression results. Through researching and collecting relevant information, it was found that the implementation of the low-carbon city pilot policy and the carbon emissions trading pilot policy overlapped in time with the smart city pilot policy, which may affect the estimation results of the model in this paper. Regarding the low-carbon city pilot policy, the "National Low-Carbon City Pilot Progress Assessment Report" published by China's Ministry of Ecology and Environment states that since 2010, China has conducted 81 national low-carbon city pilot projects in three batches. These projects have boldly explored five aspects: model innovation, system innovation, technological innovation, engineering innovation, and collaborative innovation, achieving positive results. As for the carbon emissions trading pilot policy, based on China's 12th Five-Year Plan, the carbon emissions trading pilot program was launched in 2011 in seven provinces (municipalities), including Beijing. The first carbon trading market went online in June 2013 (Shenzhen), and by June 2014, the carbon trading markets of all the pilot cities in the plan had begun operation, establishing the basic framework of China's carbon trading pilot markets. The timing for announcing emission control standards and lists of controlled enterprises across the pilot areas was concentrated in 2012–2013, with the actual opening of the carbon trading markets occurring mainly between 2013 and 2014. The implementation of the pilot policies effectively reduced carbon emissions of enterprises in the pilot regions (Li et al., 2023). To avoid interference from these

TABLE 3 PSM-DID results.

Variables	(1)	(2)	(3)
	Nearest-neighbour matching	Calliper matching (calliper = 0.04525)	Nuclear matching
	Intensity	Intensity	Intensity
DID	−0.159***	−0.158***	−0.161***
	(0.057)	(0.057)	(0.057)
Constant	4.847***	4.847***	4.847***
	(1.006)	(1.006)	(1.006)
Control Variables	YES	YES	YES
Firm Fixed Effect	YES	YES	YES
Year Fixed Effect	YES	YES	YES
City Fixed Effect	YES	YES	YES
N	17,700	17,602	16,482
R^2	0.251	0.257	0.277

TABLE 4 Control competition policy.

Variables	(1)	(2)	(3)
	Intensity	Intensity	Intensity
DID	−0.153***	−0.154***	−0.149***
	(0.057)	(0.056)	(0.057)
DID2	−0.022		−0.020
	(0.024)		(0.024)
DID3		−0.043	−0.042
		(0.027)	(0.027)
Constant	4.845***	4.854***	4.852***
	(1.008)	(1.008)	(1.009)
Control Variables	YES	YES	YES
Firm Fixed Effect	YES	YES	YES
Year Fixed Effect	YES	YES	YES
City Fixed Effect	YES	YES	YES
N	17,782	17,782	17,782
R^2	0.257	0.257	0.257

policies, dummy variables for the corresponding policies are added in the regression analysis to control for the potential impact on the estimation results.

To account for the potential impact of other climate-related policies, our regression model includes variables representing low-carbon city initiatives (DID2) and carbon emission trading schemes (DID3). Even after accounting for other policies, our key finding holds true: the smart city pilot policy is linked to a significant decrease in corporate carbon emissions, as shown in Table 4.

4.4.4 Controlling non-parallel trends and adjusting clustering standard errors

Following the methodology of Liu X. et al. (2022), we included linear, quadratic, and cubic time trend terms to capture the influence of time-dependent factors on firms. This is shown in Column 1 of Table 5. We also addressed potential issues related to standard error clustering. To address potential serial correlation, we initially clustered our data at the firm level. However, to address potential time-based variations within firms, we further clustered at the firm-year level (Column 2) and, considering the city-level implementation of the policy, at the city level (Column 3). Finally, to control for unobserved time-varying confounding factors, we included firm-year, industry-year, and city-year interaction terms in our DID model (Columns 4, 5, and 6).

The regression results consistently demonstrate a statistically significant and negative coefficient for our core explanatory variable, even after implementing various robustness checks. These checks include controlling for non-parallel trends, adjusting clustered standard errors at different levels, and incorporating fixed effects. This consistent finding strongly supports the robustness of our baseline regression results.

5 Further analysis

5.1 Heterogeneity analysis

5.1.1 Heterogeneity across high-tech and traditional industries

Companies in high-tech industries typically employ advanced technologies and production methods, characterized by strong innovation capabilities, high growth potential, and significant value (Cui and Mak, 2002). In contrast, traditional industries may rely more on resource-intensive sectors and outdated technologies, with policies potentially varying in terms of environmental and emission reduction measures for different

TABLE 5 Controlling non-parallel trends and adjusting clustering standard errors.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Controlling for non-parallel trends	Cluster to enterprise-year	Cluster to cities	Control enterprise × time	Control industry × time	Control city × time
	Intensity	Intensity	Intensity	Intensity	Intensity	Intensity
DID	−0.095**	−0.159**	−0.159**	−0.120**	−0.159***	−0.147***
	(0.046)	(0.069)	(0.063)	(0.056)	(0.057)	(0.056)
Constant	4.827***	4.847***	4.847***	90.104***	4.844***	−12.172
	(0.995)	(1.004)	(0.991)	(11.759)	(1.007)	(9.574)
Control Variables	YES	YES	YES	YES	YES	YES
Number of clusters (id)	1,368	1,368	——	1,368	1,368	1,368
Number of clusters (year)	——	13	——	——	——	——
Number of clusters (city)	——	——	260	——	——	——
Number of clusters (industry)	64	——	——	——	——	——
Firm Fixed Effect	YES	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES	YES
City Fixed Effect	YES	YES	YES	YES	YES	YES
N	17,782	17,782	17,782	17,782	17,782	17,782
R^2	0.245	0.257	0.257	0.264	0.257	0.257

industries. Given the significant differences in technological level and innovation capabilities, this paper conducts a heterogeneity analysis by categorizing companies into high-tech and traditional industries, referencing the approaches of Li and Peng (2022), to more accurately assess the impact of smart cities on corporate carbon emissions. Specifically, according to the 2012 classification guidelines for Chinese listed companies by the China Securities Regulatory Commission, companies with classification codes C25-C29, C31-C32, C34-C41, I63-I65, and M73 are defined as high-tech industry companies. After classification, there are 934 companies and 11,462 observations in the high-tech industry group, and 434 companies and 6,322 observations in the traditional industry group.

As shown in Table 6, smart city pilot policies demonstrate a statistically significant impact on carbon emission reduction within traditional industries. However, the same policies do not yield a statistically significant effect on the carbon intensity of high-tech enterprises. The primary reason for this disparity is that traditional industries are more susceptible to the direct influence of technological and management innovations associated with smart cities. Traditional industries often employ more conventional production processes and management methods, making it easier to implement and apply smart city policy measures such as intelligent production monitoring systems and energy utilization optimization technologies. In contrast, high-tech industries already possess advanced production technologies and management practices, resulting in relatively low carbon emissions. Therefore, the impact of smart city policies on their carbon emission intensity is less pronounced.

5.1.2 Heterogeneous effects across industries with varying pollution levels

Industries characterized by heavy pollution typically use energy-intensive production processes and have higher carbon emissions, whereas light industries tend to adopt cleaner production technologies resulting in relatively lower carbon emissions. Considering the significant differences in carbon emission levels, production processes, and energy consumption between them, this paper conducts a heterogeneity analysis by categorizing companies into heavy pollution and light pollution industries. This aims to more accurately evaluate the impact of smart city policies on the carbon emission intensity of different types of enterprises and to develop targeted emission reduction measures and policy support. According to the definitions of heavily polluting industries from the "Notice on Issuing the Classified Directory of Environmental Inspection Industries for Listed Companies" by the Ministry of Ecology and Environment of the People's Republic of China (formerly the Ministry of Environmental Protection of the People's Republic of China), combined with the "Industry Classification Guidelines for Listed Companies" revised by the China Securities Regulatory Commission in 2012, and following the approach of Pan et al. (2019), the codes for heavily polluting industries are specified as B06, B07, B08, B09, C17, C19, C22, C25, C26, C28, C29, C30, C31, C32, and D44. Based on this classification,

TABLE 6 Heterogeneity across high-tech and traditional industries.

Variables	High-tech industry		Traditional industries	
	(1)	(2)	(3)	(4)
	Intensity	Intensity	Intensity	Intensity
DID	−0.085	−0.083	−0.238**	−0.249**
	(0.070)	(0.069)	(0.106)	(0.104)
Constant	0.009	4.743***	0.116***	5.216***
	(0.021)	(1.280)	(0.037)	(1.686)
Control Variables	NO	YES	NO	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
City Fixed Effect	YES	YES	YES	YES
N	11,454	11,454	6,311	6,311
R^2	0.241	0.261	0.249	0.279

the study samples are divided into two groups: heavy pollution and light pollution industries, for grouped regression analysis, with the regression results shown in Table 7.

The regression results reveal that the coefficient of DID is negative in both groups of samples. The coefficient has a larger absolute value and higher significance level in the sample of heavily polluting industry enterprises. This suggests that the pilot policy has a greater and more significant impact on reducing carbon emission intensity in these industries.

The primary reason for this disparity lies in production processes in heavily polluting industries. These industries typically employ traditional energy-intensive processes with high emission levels. The measures promoted by smart city policies, such as intelligent monitoring systems and resource utilization optimization technologies, can be more directly applied to heavily polluting industries. Consequently, these policies enhance production efficiency and optimize energy use, ultimately lowering carbon emission intensity in heavily polluting industries. Conversely, lightly polluting industries, having already adopted cleaner production processes with inherently lower emissions, experience a less pronounced impact on their carbon intensity from smart city initiatives.

5.1.3 Heterogeneity across central and non-central city enterprises

This paper also conducts a heterogeneity analysis by distinguishing between central and non-central cities to examine the heterogeneous impact of smart city pilot policies on enterprise location. Central cities typically have more concentrated economic activities and population density, with a larger number and scale of enterprises. Their carbon emission intensity is significantly influenced by factors such as transportation, buildings, and industry. In practice, provincial and sub-provincial cities in China are selected as central cities. Non-central cities, on the other hand, may have different industrial structures and development models, with relatively smaller enterprise scales but unique characteristics in energy utilization and transportation organization. In practice, prefecture-level cities, excluding central cities, are chosen as the statistical standard for non-central cities.

Table 8 demonstrates that smart city pilot policies effectively curb carbon emission intensity among enterprises situated in central cities, while the same impact remains statistically insignificant in non-central cities. This divergence likely stems from the distinct characteristics inherent to central urban environments. These cities typically boast denser economic landscapes, sophisticated infrastructure, and heightened levels of technological integration. Such attributes facilitate greater accessibility and adoption of cutting-edge energy efficiency and emission reduction technologies, enabling enterprises to streamline production processes and ultimately diminish their carbon footprint.

TABLE 7 Heterogeneous effects across industries with varying pollution levels.

Variables	Heavily polluting industries		Lightly polluting industries	
	(1)	(2)	(3)	(4)
	Intensity	Intensity	Intensity	Intensity
DID	−0.176**	−0.181**	−0.146*	−0.154*
	(0.087)	(0.084)	(0.083)	(0.081)
Constant	0.070**	6.302***	0.034	4.508***
	(0.028)	(1.815)	(0.026)	(1.222)
Control Variables	NO	YES	NO	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
City Fixed Effect	YES	YES	YES	YES
N	8,056	8,056	9,709	9,709
R^2	0.234	0.264	0.254	0.271

TABLE 8 Heterogeneity across central and non-central city enterprises.

Variables	Central cities		Non-central cities	
	(1)	(2)	(3)	(4)
	Intensity	Intensity	Intensity	Intensity
DID	−0.279 ***	−0.295 ***	−0.060	−0.055
	(0.105)	(0.105)	(0.067)	(0.066)
Constant	0.095 **	3.719 **	0.022	5.450 ***
	(0.038)	(1.472)	(0.019)	(1.334)
Control Variables	NO	YES	NO	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
City Fixed Effect	YES	YES	YES	YES
N	7,371	7,371	10,406	10,406
R^2	0.258	0.282	0.219	0.244

However, in non-central cities, factors such as relatively simple industrial structures, smaller enterprise sizes and lagging technological levels may limit the impact of pilot policies. This condition could be attributed to a lack of advanced technological support, insufficient resource allocation and limitations in policy implementation.

5.2 Mechanism analysis

Based on the above analysis, the pilot policies primarily operate through these three channels: promoting green technology innovation, increasing external market attention and increasing government subsidies. In mechanism analysis, most scholars use the mediating effect model. However, this method may lead to unreliable results due to endogeneity bias and other issues. This study draws on the research method of Liu and Mao (2019) to explore the impact of mechanism variables on the treatment effect of pilot policies from the perspective of moderating effect.

Findings from Table 9 demonstrate that smart city pilot policies effectively reduce enterprise carbon emission intensity, with a statistically significant positive coefficient for the policy dummy variable (DID) at the 1% level. This effect is achieved by promoting green technology innovation. The pilot policy is committed to integrating advanced information technology, intelligent systems and sustainable development concepts to improve urban operation efficiency, resource utilization efficiency and environmental quality. As smart cities develop, enterprises are increasingly driven to invest in green technology innovation. This trend is propelled by escalating environmental pressures and stricter carbon emission regulations. By introducing efficient energy-saving technologies and clean production processes, green technology innovation can effectively reduce energy consumption and carbon emissions during the production process, achieving cleaner and low-carbon production. Green technology innovation not only addresses environmental pressures but also offers tangible operational benefits. By optimizing resource utilization efficiency, it directly contributes to reducing the carbon footprint of enterprise production processes. Thus, Hypothesis 2 is verified.

Regression results presented in columns 2 and 3 of Table 6 indicate that smart city pilot policies demonstrably increase both investor and media attention towards enterprises. The pilot policy, by guiding external market attention, especially the supervisory pressure from investors and the media, provides an important external driving force for enterprises to reduce carbon emission intensity. The attention of investors and the media to corporate environmental protection and sustainable development encourages enterprises to pay more attention to emission reduction. Firstly, the

TABLE 9 Mechanism analysis.

Variables	Green technology innovation	External attention		Government subsidies	
	(1)	(2)	(3)	(4)	(5)
	Innovation	Investor-attention	Media-attention	Ex-ante subsidy	Ex-post subsidy
DID	0.549***	0.243***	0.249**	0.353***	−0.187
	(0.022)	(0.058)	(0.122)	(0.132)	(0.272)
Constant	−0.523	1.978**	2.436	14.248***	22.958***
	(0.442)	(0.794)	(2.279)	(2.141)	(5.027)
Control Variables	YES	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES	YES
City Fixed Effect	YES	YES	YES	YES	YES
N	17,782	17,782	17,782	17,782	17,782
R^2	0.702	0.889	0.774	0.529	0.629

increased attention of investors has improved the environmental transparency and sustainable development evaluation standards of enterprises. To maintain good investor relations and attract more investment, enterprises tend to take more proactive emission reduction measures to enhance their environmental image and investment attractiveness. Secondly, continuous media attention can increase the exposure and public awareness of enterprises. It often increases public attention to corporate environmental actions, thereby shaping the social image and reputation of enterprises. Under the pressure of public attention, enterprises will take more proactive emission reduction measures to maintain their image and reputation. Therefore, the smart city pilot policy effectively reduces the carbon emission intensity of enterprises by increasing the attention of investors and the media, encouraging enterprises to take more active emission reduction measures. Thus, Hypothesis 3 is verified.

The fourth column's DID coefficient demonstrates that the smart city pilot policy, by increasing government subsidies, effectively reduces enterprise carbon emission intensity. These subsidies, a key fiscal policy tool, lower the cost for businesses investing in green technology adoption and energy-efficient, emission-reducing equipment. This financial incentive stimulates companies to proactively pursue emission reduction measures. In addition, government subsidies can be used to support enterprises in carrying out carbon emission monitoring and reporting. By providing subsidies or rewards to enterprises, they are encouraged to establish and implement carbon emission monitoring and reporting systems, strengthen the supervision and management of enterprise carbon emissions, and encourage enterprises to take proactive emission reduction measures to reduce carbon emission intensity. Therefore, the smart city pilot policy effectively promotes enterprise emission reduction actions and reduces the carbon emission intensity of enterprises by increasing government subsidies. Thus, Hypothesis 4 is verified.

6 Conclusion and policy implications

This paper uses data from Chinese A-share listed companies from 2009 to 2021 as a research sample to deeply analyze the theoretical logic and underlying mechanisms by which smart city construction affects corporate carbon emission intensity. Based on this analysis, a multi-period differences-in-differences model is constructed to empirically examine the impact of smart city pilot policies on corporate carbon emission intensity. The study finds that: ① Smart city pilot policies significantly reduce the carbon emission intensity of enterprises in pilot cities, and this conclusion remains robust after controlling for competitive policies and non-parallel trends. ② The impact of smart city pilot policies on corporate carbon emission intensity varies across different industries and cities, with more pronounced effects observed in central cities, traditional industries, and heavily polluting industries. ③ Further mechanism analysis indicates that smart city pilot policies reduce corporate carbon emission intensity through three mechanisms: promoting green technological innovation, increasing external market attention, and enhancing pre-emptive government subsidies.

Based on the above conclusions, this paper offers the following policy implications: First, continued investment and development of smart cities will create significant opportunities for reducing corporate carbon emissions. By providing targeted support to enterprises and fostering a sustainable ecosystem, smart city initiatives can effectively lower carbon intensity, encouraging businesses to embrace low-carbon development models and promoting a virtuous cycle of economic growth and environmental protection. Local governments should prioritize the development of robust smart infrastructure and promote data sharing and transparency. Establishing cross-sector and cross-industry data sharing mechanisms will be crucial. This will facilitate information flow and cross-border collaboration, ultimately providing enterprises with enhanced opportunities and support for carbon emission reduction initiatives. At the same time, policy guidance and incentive measures should be strengthened to encourage enterprises to adopt more environmentally friendly and energy-saving production methods, cooperation and exchanges between smart city construction and enterprises should be enhanced, and enterprises should be encouraged to participate in smart city planning, construction and operation to jointly promote low-carbon development.

Second, the government should implement differentiated smart city pilot policies tailored to enterprises across various industries and cities, taking into full account their unique characteristics and needs to promote green and low-carbon development. For central cities, it is essential to continue implementing smart city pilot policies, focusing on enhancing the intelligence level of environmental protection public facilities, leveraging economies of scale, and reducing the costs of emission reduction and carbon abatement for enterprises. In contrast, for non-central cities, it is crucial to promptly adjust smart city pilot construction plans, considering challenges such as relatively lagging economic development, inadequate information infrastructure, and weak green transition capabilities. Emphasis should be placed on leveraging the networked and shared characteristics of modern information systems to actively develop green and intelligent network service facilities with cross-regional service functions, thereby reducing the costs of green development in non-central cities. For traditional and heavily polluting industries, where pilot policies have shown significant effects on corporate carbon reduction, it is important to continue implementing these policies while emphasizing the application and secondary innovation of modern information technology in these sectors. Through intelligent construction, enterprises can enhance their green technology innovation capabilities. For high-tech and low-pollution industries, where the potential for carbon reduction through pilot policies is limited, it is necessary to strengthen industrial technology exchanges. This approach can drive the update of green production technologies and improve energy-saving and emission-reduction efficiency in traditional and heavily polluting industries through the development of high-tech industries.

Third, enhancing smart city development and effectively reducing the carbon emission intensity of enterprises require collaborative efforts from the government, businesses, investors, and the media. The government should establish a comprehensive policy framework that includes specific incentives and guiding policies to encourage increased investment in green technology

innovation by enterprises. Various government subsidy types play distinct roles in promoting corporate carbon reduction within low-carbon city initiatives. Therefore, the government should strengthen the implementation of pre-subsidy policies, rigorously monitor and evaluate the use of subsidy funds, and further amplify the impact of pilot programs on corporate carbon reduction. Additionally, the government should utilize modern communication tools and social media platforms to promote and guide sustainable development concepts, such as green development and lifestyles, facilitating the transition of businesses to green production methods and encouraging residents to adopt greener lifestyles. The government must fulfill its environmental oversight responsibilities by enhancing environmental protection laws and regulations, strengthening the supervision of carbon reduction efforts in polluting enterprises, and reducing the carbon emission intensity of businesses.

Fourth, the effectiveness of smart city pilot policies depends on the specific developmental stage and politico-economic conditions of a region. As a government-led initiative, smart city pilot policies represent a significant urban intelligence construction project. Given that China is at a critical juncture as a developing country aiming for carbon peaking, it faces the dual challenges of carbon reduction and economic development. Successfully achieving a green and low-carbon economic transition requires fully leveraging the government's regulatory role in the economy while respecting market principles, technological advancement, and sustainable development. The study's conclusions and policy recommendations are applicable to other developing countries at a similar stage of development, facing the dual challenges of economic growth and carbon peaking. Additionally, ensuring the scientific formulation and effective implementation of smart city pilot policies necessitates a strong and incorruptible government.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

WP: Conceptualization, Funding acquisition, Investigation, Writing–original draft, Writing–review and editing. JL: Data curation, Investigation, Validation, Writing–original draft, Writing–review and editing. XR: Conceptualization, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by National Social Science Foundation, grant number 21BJL078.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2024.1457801/full#supplementary-material

References

Ahern, K. R., and Sosyura, D. (2015). Rumor has it: sensationalism in financial media. *Rev. Financial Stud.* 28, 2050–2093. doi:10.1093/rfs/hhv006

Cai, X., Lu, Y., Wu, M., and Yu, L. (2016). Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. *J. Dev. Econ.* 123, 73–85. doi:10.1016/j.jdeveco.2016.08.003

Caragliu, A., and Del Bo, C. F. (2019). Smart innovative cities: the impact of smart city policies on urban innovation. *Technol. Forecast. Soc. Change* 142, 373–383. doi:10.1016/j.techfore.2018.07.022

Chapple, L., Clarkson, P. M., and Gold, D. L. (2013). The cost of carbon: capital market effects of the proposed emission trading scheme (ETS). *Abacus* 49, 1–33. doi:10.1111/abac.12006

Chen, H., Wang, R., Liu, X., Du, Y., and Yang, Y. (2023). Monitoring the enterprise carbon emissions using electricity big data: a case study of Beijing. *J. Clean. Prod.* 396, 136427. doi:10.1016/j.jclepro.2023.136427

Chen, Q., Jiao, J., Wang, Y., Mai, Z., Ren, J., He, S., et al. (2020). Egr-1 mediates low-dose arecoline induced human oral mucosa fibroblast proliferation via transactivation of Wnt5a expression. *Nankai Econ. Stud.* 21 (6), 80–100. doi:10.1186/s12860-020-00325-7

Chen, S., Mao, H., and Sun, J. (2022). Low-carbon city construction and corporate carbon reduction performance: evidence from a quasi-natural experiment in China. *J. Bus. Ethics* 180, 125–143. doi:10.1007/s10551-021-04886-1

Chu, Z., Cheng, M., and Yu, N. N. (2021). A smart city is a less polluted city. *Technol. Forecast. Soc. Change* 172, 121037. doi:10.1016/j.techfore.2021.121037

Cui, H., and Mak, Y. T. (2002). The relationship between managerial ownership and firm performance in high R&D firms. *J. Corp. Finance* 8 (4), 313–336. doi:10.1016/s0929-1199(01)00047-5

Cui, J., Lapan, H., and Moschini, G. (2016). Productivity, export, and environmental performance: air pollutants in the United States. *Am. J. Agric. Econ.* 98, 447–467. doi:10.1093/ajae/aav066

Dahlmann, F., Branicki, L., and Brammer, S. (2019). Managing carbon aspirations: the influence of corporate climate change targets on environmental performance. *J. Bus. Ethics* 158, 1–24. doi:10.1007/s10551-017-3731-z

Feng, Y., Gao, Y., Meng, X., Shi, J., Shi, K., Hu, S., et al. (2023). The impacts of casual environmental regulation on carbon intensity in China: dual mediating pathways of energy low-carbon reconstitution and industrial structure upgrading. *Environ. Res.* 238, 117289. doi:10.1016/j.envres.2023.117289

Feng, Y., Sun, M., Pan, Y., and Zhang, C. (2024). Fostering inclusive green growth in China: identifying the impact of the regional integration strategy of Yangtze River economic belt. *J. Environ. Manag.* 358, 120952. doi:10.1016/j.jenvman.2024.120952

Gao, K., and Yuan, Y. (2022). Is the sky of smart city bluer? Evidence from satellite monitoring data. *J. Environ. Manag.* 317, 115483. doi:10.1016/j.jenvman.2022.115483

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *J. Econ.* 225, 254–277. doi:10.1016/j.jeconom.2021.03.014

Guo, C., Wang, Y., Hu, Y., Wu, Y., and Lai, X. (2024). Does smart city policy improve corporate green technology innovation? Evidence from Chinese listed companies. *J. Environ. Plan. Manag.* 67 (6), 1182–1211. doi:10.1080/09640568.2022.2157708

Guo, Q., Wang, Y., and Dong, X. (2022). Effects of smart city construction on energy saving and CO_2 emission reduction: evidence from China. *Appl. Energy* 313, 118879. doi:10.1016/j.apenergy.2022.118879

Han, L., Chen, S., and Liang, L. L. (2021). Digital economy, innovation environment and urban innovation capabilities. *Sci. Res. Manag.* 42 (4), 35–45. doi:10.19571/j.cnki.1000-2995.2021.04.004

Haney, A. B. (2017). Threat interpretation and innovation in the context of climate change: an ethical perspective. *J. Bus. Ethics* 143, 261–276. doi:10.1007/s10551-015-2591-7

Imbens, G. W., and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *J. Econ. Literature* 47 (1), 5–86. doi:10.1257/jel.47.1.5

Ivanov, D., Dolgui, A., and Sokolov, B. (2022). Cloud supply chain: integrating industry 4.0 and digital platforms in the supply chain-as-a-service. *Logist. Transp. Rev.* 160, 102676. doi:10.1016/j.tre.2022.102676

Jiang, H., Jiang, P., Wang, D., and Wu, J. (2021). Can smart city construction facilitate green total factor productivity? A quasi-natural experiment based on China's pilot smart city. *Sustain. Cities Soc.* 69, 102809. doi:10.1016/j.scs.2021.102809

Kumarasiri, J. (2017). Stakeholder pressure on carbon emissions: strategies and the use of management accounting. *Australas. J. Environ. Manag.* 24, 339–354. doi:10.1080/14486563.2017.1350210

Li, B. (2022). Effective energy utilization through economic development for sustainable management in smart cities. *Energy Rep.* 8, 4975–4987. doi:10.1016/j.egyr.2022.02.303

Li, W., and Peng, H. (2022). The situation and reflections on technological innovation investment of Chinese listed companies: also on the enhancement of the main position of enterprise innovation. *Quantitative Technol. Econ. Res.* (06), 100–119. doi:10.13653/j.cnki.jqte.2022.06.001

Li, Z., Che, S., and Wang, J. (2023). Analysis of the policy effectiveness of the implementation of China's carbon emissions trading pilot: new evidence from PSM-DID and SCM. *Manag. Rev.* 35 (12), 308–318. doi:10.14120/j.cnki.cn11-5057/f.2023.12.011

Liu, K., Tao, Y., Wu, Y., and Wang, C. (2020). How does ecological civilization construction affect carbon emission intensity? Evidence from Chinese provinces' panel data. *Chin. J. Popul. Resour. Environ.* 18, 97–102. doi:10.1016/j.cjpre.2019.10.002

Liu, L., Wang, Y., and Xu, Y. (2022). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. *Am. J. Political Sci.* 68, 160–176. doi:10.1111/ajps.12723

Liu, X., Cifuentes-Faura, J., Zhao, S., and Wang, L. (2023). Government environmental attention and carbon emissions governance: firm-level evidence from China. *Econ. Analysis Policy* 80, 121–142. doi:10.1016/j.eap.2023.07.016

Liu, X., Zhang, K., and Ren, Y. (2022). Does climate warming affect labour productivity in emerging economies? Evidence from Chinese-listed firms. *Appl. Econ.* 55 (24), 2801–2814. doi:10.1080/00036846.2022.2106033

Liu, Y., and Mao, J. (2019). How do tax incentives affect investment and productivity? Firm-level evidence from China. *Am. Econ. J. Econ. Policy* 11, 261–291. doi:10.1257/pol.20170478

Mahmood, H., Alkhateeb, T. T. Y., and Furqan, M. (2020). Industrialization, urbanization and CO_2 emissions in Saudi Arabia: asymmetry analysis. *Energy Rep.* 6, 1553–1560. doi:10.1016/j.egyr.2020.06.004

Mora, C., Spirandelli, D., Franklin, E. C., Lynham, J., Kantar, M. B., Miles, W., et al. (2018). Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions. *Nat. Clim. Change* 8, 1062–1071. doi:10.1038/s41558-018-0315-6

Nicolas, C., Kim, J., and Chi, S. (2021). Natural language processing-based characterization of top-down communication in smart cities for enhancing citizen alignment. *Sustain. Cities Soc.* 66, 102674. doi:10.1016/j.scs.2020.102674

Niu, S., Chen, Y., Zhang, R., and Feng, Y. (2023). How does the air pollution prevention and control action plan affect sulfur dioxide intensity in China? *Front. Public Health* 11, 1119710. doi:10.3389/fpubh.2023.1119710

Pan, A., Liu, X., Qiu, J., and Shen, Y. (2019). Can green mergers and acquisitions under media pressure promote substantial transformation of heavily polluting enterprises? *China Ind. Econ.* (02), 174–192. doi:10.19581/j.cnki.ciejournal.20190131.005

Peng, H., and Liu, Y. (2018). How government subsidies promote the growth of entrepreneurial companies in clean energy industry: an empirical study in China. *J. Clean. Prod.* 188, 508–520. doi:10.1016/j.jclepro.2018.03.126

Richter, P. M., and Schiersch, A. (2017). CO_2 emission intensity and exporting: evidence from firm-level data. *Eur. Econ. Rev.* 98, 373–391. doi:10.1016/j.euroecorev.2017.07.011

Sadler, T. R. (2016). Institutional pressures and organizational characteristics: the case of polluting emissions and the toxics release inventory. *J. Interdiscip. Econ.* 28, 1–23. doi:10.1177/0260107915609826

Shen, Q., Wu, R., Pan, Y., and Feng, Y. (2023). The effectiveness of smart city policy on pollution reduction in China: new evidence from a quasi-natural experiment. *Environ. Sci. Pollut. Res.* 30, 52841–52857. doi:10.1007/s11356-023-26010-8

Shen, Q., Wu, R., Pan, Y., and Feng, Y. (2024). Explaining and modeling the impacts of inclusive finance on CO_2 emissions in China integrated the intermediary role of energy poverty. *Humanit. Soc. Sci. Commun.* 11, 82–19. doi:10.1057/s41599-023-02595-w

Wang, C., Wang, W., and Huang, R. (2017). Supply chain enterprise operations and government carbon tax decisions considering carbon emissions. *J. Clean. Prod.* 152, 271–280. doi:10.1016/j.jclepro.2017.03.051

Wang, F., Sun, J., and Liu, Y. S. (2019). Institutional pressure, ultimate ownership, and corporate carbon reduction engagement: evidence from China. *J. Bus. Res.* 104, 14–26. doi:10.1016/j.jbusres.2019.07.003

Wang, H., Wang, S., Miao, Z., et al. (2016). Heterogeneity threshold effect of R&D investment on green innovation efficiency: based on Chinese high-tech industries. *Sci. Res. Manag.* 37 (2), 63–71. doi:10.19571/j.cnki.1000-2995.2016.02.008

Wang, L., Chen, Y., Ramsey, T. S., and Hewings, G. J. (2021). Will researching digital technology really empower green development? *Technol. Soc.* 66, 101638. doi:10.1016/j.techsoc.2021.101638

Wu, S. (2022). Smart cities and urban household carbon emissions: a perspective on smart city development policy in China. *J. Clean. Prod.* 373, 133877. doi:10.1016/j.jclepro.2022.133877

Xie, Z., Wu, R., and Wang, S. (2021). How technological progress affects the carbon emission efficiency? Evidence from national panel quantile regression. *J. Clean. Prod.* 307, 127133. doi:10.1016/j.jclepro.2021.127133

Xin, B., and Qu, Y. (2019). Effects of smart city policies on green total factor productivity: evidence from a quasi-natural experiment in China. *Int. J. Environ. Res. Public Health* 16, 2396. doi:10.3390/ijerph16132396

Xu, A., Wang, W., and Zhu, Y. (2023). Does smart city pilot policy reduce CO_2 emissions from industrial firms? Insights from China. *J. Innovation Knowl.* 8 (3), 100367. doi:10.1016/j.jik.2023.100367

Yang, S., Jahanger, A., and Usman, M. (2024). Examining the influence of green innovations in industrial enterprises on China's smart city development. *Technol. Forecast. Soc. Change* 199, 123031. doi:10.1016/j.techfore.2023.123031

Yao, T., Huang, Z., and Zhao, W. (2020). Are smart cities more ecologically efficient? Evidence from China. *Sustain. Cities Soc.* 60, 102008. doi:10.1016/j.scs.2019.102008

Yu, F., Han, F., and Cui, Z. (2015). Reducing carbon emissions through industrial symbiosis: a case study of a large enterprise group in China. *J. Clean. Prod.* 103, 811–818. doi:10.1016/j.jclepro.2014.05.038

Zawieska, J., and Pieriegud, J. (2018). Smart city as a tool for sustainable mobility and transport decarbonisation. *Transp. Policy* 63, 39–50. doi:10.1016/j.tranpol.2017.11.004

Zhan, Y., and Li, S. (2022). Smart city construction, entrepreneurial vitality and high-quality economic development: analysis based on the GTFP perspective. *J. Finance Econ.* 48 (1), 4–18. doi:10.16538/j.cnki.jfe.20211015.101

frontiers | Frontiers in Environmental Engineering

TYPE Original Research
PUBLISHED 31 March 2025
DOI 10.3389/fenve.2025.1488965

OPEN ACCESS

EDITED BY
Sayali Sandbhor,
Symbiosis International University, India

REVIEWED BY
Yeunook Bae,
Northwestern Medicine, United States
Asude Hanedar,
Namik Kemal University, Türkiye

*CORRESPONDENCE
Pradeep Kurup,
pradeep_Kurup@uml.edu
Mohammad Arif Ul Alam,
mohammadariful_alam@uml.edu

RECEIVED 31 August 2024
ACCEPTED 27 February 2025
PUBLISHED 31 March 2025

CITATION
Anaadumba R, Bozkurt Y, Sullivan C, Pagare M, Kurup P, Liu B and Alam MAU (2025) Graph neural network-based water contamination detection from community housing information.
Front. Environ. Eng. 4:1488965.
doi: 10.3389/fenve.2025.1488965

Graph neural network-based water contamination detection from community housing information

Raphael Anaadumba, Yigit Bozkurt, Connor Sullivan, Madhavi Pagare, Pradeep Kurup*, Benyuan Liu and Mohammad Arif Ul Alam*

University of Massachusetts Lowell, Lowell, MA, United States

Introduction: Detecting water contamination in community housing is crucial for protecting public health. Early detection enables timely action to prevent waterborne diseases and ensures equitable access to safe drinking water. Traditional methods recommended by the Environmental Protection Agency (EPA) rely on collecting water samples and conducting lab tests, which can be both time-consuming and costly.

Methods: To address these limitations, this study introduces a Graph Attention Network (GAT) to predict lead contamination in drinking water. The GAT model leverages publicly available municipal records and housing information to model interactions between homes and identify contamination patterns. Each house is represented as a node, and relationships between nodes are analyzed to provide a clearer understanding of contamination risks within the community.

Results: Using data from Flint, Michigan, the model demonstrated higher performance compared to traditional methods. Specifically, the GAT achieved an accuracy of 0.80, precision of 0.71, and recall of 0.93, outperforming XGBoost, a classical machine learning algorithm, which had an accuracy of 0.70, precision of 0.66, and recall of 0.67.

Discussion: In addition to its predictive capabilities, the GAT model identifies key factors contributing to lead contamination, enabling more precise targeting of at-risk areas. This approach offers a practical tool for policymakers and public health officials to assess and mitigate contamination risks, ultimately improving community health and safety.

KEYWORDS

water contamination, public health, graph attention network (GAT), environmental hazards, flint Michigan

1 Introduction

Detecting contamination in drinking water is critical for safeguarding public health, as it prevents exposure to harmful pollutants and ensures access to safe drinking water. In the United States, nearly one-fifth of the population, approximately 63 million people, have been exposed to potentially unsafe water multiple times over the past decade (News21, 2023). An investigation by the Environmental Protection Agency (EPA) reported over

680,000 water quality violations, affecting communities ranging from rural Central California to urban New York City. This widespread issue underscores the need for effective detection methods that address both the severity and scale of contamination. Health risks associated with lead in drinking water are well-documented, with even minimal concentrations posing significant concerns. Although the EPA has established an action level of 15 ppb for lead (Pb), no concentration is considered entirely safe because lead is non-biodegradable and can accumulate in the food chain (Doré et al., 2020; Han et al., 2020; Sawan et al., 2020). Prolonged exposure can harm the brain, kidneys, and other organs (Vlachou et al., 2020; Martin and Griswold, 2009; Trueman et al., 2016).

Lead service lines (LSLs) are a primary source of contamination, contributing 50%–75% of the lead mass in drinking water, followed by premise piping (20%–35%) and faucets (1%–3%) (Sandvig et al., 2008). In the U.S., an estimated 6.1 to 10.2 million LSLs supply water to approximately 15–22 million people, accounting for about 7% of community water system consumers (Cornwell et al., 2016; Hensley et al., 2021). Addressing LSLs is critical for reducing lead exposure risks and safeguarding public health.

Various methods exist for detecting LSLs, including record screening, visual inspections, water quality testing (e.g., EPA methods 200.8 and 200.9), excavation, and advanced techniques like cumulative lead sampling devices and acoustic wave technology (Hensley et al., 2021). More recently, predictive data analysis, particularly machine learning, has emerged as a cost-effective alternative. Machine learning enables researchers to analyze patterns in existing data, identifying high-risk areas without relying on expensive sampling or specialized equipment. By focusing on areas most at risk, these models support more efficient mitigation efforts and help reduce lead exposure in drinking water.

The extent of lead release from LSLs depends on a combination of factors, including water chemistry, pipe scale composition, and environmental conditions (Pasteris et al., 2021; García-Timermans et al., 2023). Tools such as Raman spectroscopy have been used to analyze the stability of lead-pipe scales under changing treatment conditions, providing insights into how factors like pH adjustments, disinfectants, and orthophosphate dosing affect lead release (Pasteris et al., 2021). Similarly, pilot-scale water distribution studies have shown that biofilm formation, pipe aging, and transport conditions also influence water quality (García-Timermans et al., 2023). To mitigate these factors and manage lead release, effective corrosion control strategies are essential.

Corrosion control strategies play an important role in managing lead release from service lines. Zinc orthophosphate has been effective in reducing nitrate-induced lead corrosion, particularly in systems with high nitrate levels (Lopez et al., 2024). However, challenges such as aluminum accumulation in pipe scales can delay the performance of phosphate inhibitors, indicating the need for adaptive strategies (Li et al., 2020). The composition of pipe scales influences lead dynamics: iron-rich scales accelerate the oxidation of Pb(II) to Bae et al. (2020b), while manganese facilitates the oxidation of lead carbonate, stabilizing PbO formation in chlorinated systems (Pan et al., 2019). Interactions with other metals, such as chromium, add further complexity. Chromium release from pipe scales is influenced by anion concentrations; sulfate and chloride promote release under certain conditions (Devine et al., 2024; Ni et al., 2024; Bae et al., 2020a). Orthophosphate can stabilize lead release by forming calcium–lead–phosphorus solids, but trade-offs like calcium phosphate precipitation illustrate the complexities of balancing corrosion control and water quality (Bae et al., 2020c; Devine et al., 2024). These findings highlight the intricate interplay among water chemistry, pipe materials, and environmental factors in determining lead behavior in drinking water systems. Understanding these dynamics emphasizes the need for predictive tools to effectively address lead contamination.

Machine learning offers an approach to complement traditional water quality assessments by modeling interactions between water chemistry, pipe scales, and environmental factors. These models provide insights that help identify areas of elevated contamination risk and guide targeted mitigation strategies. Unlike traditional methods that often require extensive infrastructure interventions, machine learning leverages spatial and temporal data to improve the accuracy of risk predictions. This approach addresses challenges such as variable lead levels and incomplete service line data, enabling more efficient resource allocation and decision-making.

Recent studies have applied machine learning models to address lead contamination. Random Forest models achieved cross-validation scores of 0.88 for Massachusetts and 0.78 for California (Lobo et al., 2022). In Pittsburgh, precise predictions (over 90%) were made for only 13% of customers, suggesting that unnecessary excavations could be reduced by improving short-term replacement decisions (Hajiseyedjavadi et al., 2022). Incorporating field observations of tap water materials further improved prediction accuracy to 94% when integrated into models (Blackhurst, 2021). Gradient Boosting models predicted high lead levels (over 15 ppb) in tap water for the Pittsburgh Water and Sewer Authority, achieving an AUC score of 71.6% (Hajiseyedjavadi et al., 2022). Similarly, Support Vector Machines identified lead service lines with an average accuracy of 90% (Gurewitsch, 2019). In Flint, Michigan, an XGBoost model identified 1,000 homes most likely to exceed the EPA's action level of 15 ppb, even without direct test results (Chojnacki et al., 2017).

Despite these advancements, challenges remain. Imbalanced datasets can bias model training, and incomplete or inaccurate service line data compromise prediction reliability. The variability of lead levels adds complexity, and many models focus exclusively on lead results without incorporating spatial factors such as proximity to contamination sources. Addressing these limitations is crucial for improving the accuracy and broader applicability of machine learning models.

Machine learning applications for lead contamination have been explored across diverse regions. In Flint, Michigan, researchers developed predictive models combining residential water test data with infrastructure information, though these efforts faced computational and data limitations (Abernethy et al., 2016). In Saudi Arabia, techniques like Nonlinear Autoregressive Neural Networks and Long Short-Term Memory networks were used to predict the Water Quality Index, but the findings were specific to that region (Aldhyani et al., 2020). In Chicago, Illinois, models such as Random Forest, logistic regression, and support vector machines assessed lead poisoning risks in children, often focusing on lead paint rather than waterborne contamination (Potash et al., 2015). In Pittsburgh, Pennsylvania, Support Vector Machine and Random

Forest models assessed lead service line risks in residential areas, but their performance was limited by incomplete data and a narrow focus on factors like housing age and spatial characteristics (Hajiseyedjavadi et al., 2020; Gurewitsch, 2019). Similarly, in California and Massachusetts, Random Forest models identified high-risk areas for lead in school drinking water, relying heavily on publicly available data without establishing causality (Lobo et al., 2022).

Given the complexity of environmental data and the limitations of existing models, advanced methods are required to address these challenges. This study applies GATs to lead contamination risk assessment, leveraging their ability to capture geographic relationships and complex interactions. Previous research on Graph Neural Networks (GNN) has shown their effectiveness in environmental and water management tasks. For example, Graph Convolutional Recurrent Neural Networks (GCRNN) have been applied to water demand forecasting, capturing spatial and temporal dependencies (Zanfei et al., 2022). Similarly, GNN have been used in groundwater level prediction in British Columbia, Canada, by representing wells as graph nodes and learning spatial relationships through a self-adaptive adjacency matrix (Bai and Tahmasebi, 2023). In river networks, GATs combined with spatiotemporal fusion have modeled spatial dependencies among nodes and temporal dynamics (Lin et al., 2022). These applications suggest that GATs have the potential to improve predictions of lead contamination in drinking water.

While lead service lines remain a primary source of contamination, this study focuses on developing predictive tools to assess contamination risks rather than conducting direct experimentation on service lines. By utilizing GATs, this research aims to address the limitations of existing machine learning models and enhance risk assessment accuracy. The study contributes to the field by providing a robust framework for contamination prediction, integrating spatial dependencies and environmental factors. To provide a structured view of the implemented approach, Figure 1 summarizes the key steps involved in the water contamination detection framework.

2 Materials and methods

2.1 Data collection

Our study is centered on the city of Flint, Michigan, which has been facing a significant crisis due to lead contamination in its water supply (Michigan.gov, 2023). Our initial dataset was derived from the water testing services in Flint, in collaboration with the Michigan Department of Environmental Quality (Michigan.gov, 2023). This dataset spans January through December 2016, capturing household-level water sampling efforts undertaken after the Flint water crisis. The time frame is critical, as it reflects lead contamination trends during early remediation measures, including the introduction of corrosion control. This dataset comprised approximately 14,000 records, predominantly from the year 2016 as mentioned earlier. However, the limited diversity of features within this dataset rendered it inadequate for machine learning analysis.

To address this limitation, we utilized web scraping techniques to augment our dataset with additional property information from the Zillow online real estate database (www.zillow.com) (Sarr and Diallo, 2018; Zillow, 2023). The web scraping process was implemented using Python, which provided flexibility for data retrieval and processing. The approach used libraries like `googlesearch-python` to programmatically retrieve URLs via Google searches, `requests` and `json` for HTTP requests and API responses, and `pandas` for handling and storing the data. Additional utilities, such as `io` and `time`, supported auxiliary tasks. The scraping process began by identifying relevant links using `googlesearch-python`, followed by parsing HTML content and integrating with RapidAPI to extract structured data about properties. Key attributes, such as longitude and latitude, building size, year built, and condition, were systematically stored in Comma-Separated Values (CSV) format for downstream analysis. To ensure the validity and accuracy of the web-scraped data, we cross-referenced residential addresses and features (e.g., year built) with the Michigan Department of Environmental dataset and Arc GIS parcel records. This validation step confirmed the consistency of the scraped data and ensured its reliability for further analysis. The web scraping process retrieved approximately 1,070 property records. Following the cleaning and validation steps described, 154 records were excluded due to missing over 80% of key features, such as longitude and latitude, building size, year built, and condition or due to a lack of alignment with other sources. This left 916 records that met the inclusion criteria. These excluded records were largely incomplete and unlikely to impact the findings or the model's ability to generalize contamination patterns. This cleaning and validation process was applied uniformly across the dataset to maintain integrity and avoid systematic bias. Figure 2 illustrates our entire data source.

Furthermore, we enriched our dataset with Arc GIS parcel record data from the City of Flint office, incorporating critical property details such as location and valuation to enhance the depth of our analysis. Any inconsistencies or unmatched data were omitted to maintain data integrity. Additionally, we examined seasonal variations in lead concentrations to determine whether temperature fluctuations or changes in water usage patterns influenced contamination levels. A seasonal breakdown of lead levels showed minor variations, with slightly higher concentrations in summer and fall and lower levels in winter. These trends suggest that increased corrosion in warmer months and shifts in water demand may have contributed to fluctuations. However, the differences were not large enough to indicate strong seasonal dependence.

2.2 Geographic information system mapping in flint

To provide a visual representation of our study area, Figure 3 presents a GIS map indicating the houses in Flint that were sampled for these prediction studies. Such a map offers a spatial understanding of the scope of our research and the distribution of samples across the city. The data for our study was carefully sampled homes from the initial dataset of 14,000 records. Homes with higher lead concentrations were prioritized to train the model effectively in identifying contamination patterns. This approach ensured that the dataset was representative of homes across various lead contamination levels, as categorized in Table 1.

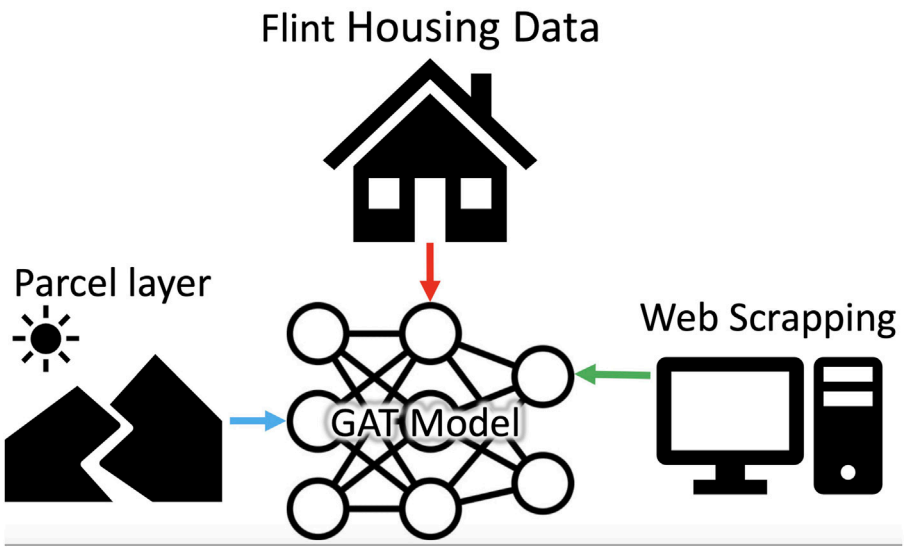


FIGURE 1
Gnn development with multiple data sources.

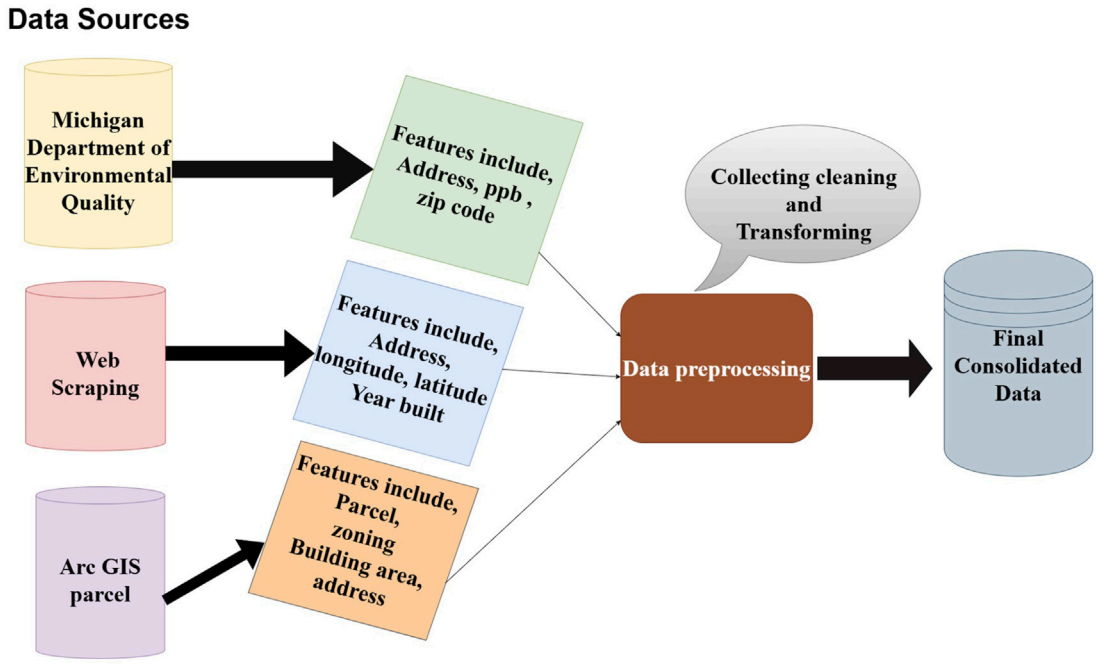


FIGURE 2
Workflow for data consolidation from multiple sources for lead contamination analysis.

2.3 Data curation

To focus on houses affected by varying lead levels, we adopted a labeling scheme aligned with EPA standards (Agency, 2023). Houses with lead levels below 5 ppb were categorized as Level 0; those with levels from five to less than 10 ppb as Level 1; from 10 to less than 15 ppb as Level 2; and houses with lead levels of 15 ppb or higher as Level 3 Table 2. These classification thresholds reflect real-world standards, enabling the model to detect and differentiate multiple levels of lead contamination severity. Recognizing that no level of lead is entirely safe, this scheme emphasizes early detection and intervention. By identifying contamination even at minimal levels, the model aims to provide actionable insights to mitigate risks and facilitate timely responses.

Out of the initial 14,000 records, only 916 homes met the inclusion criteria for our analysis. Specifically, each included

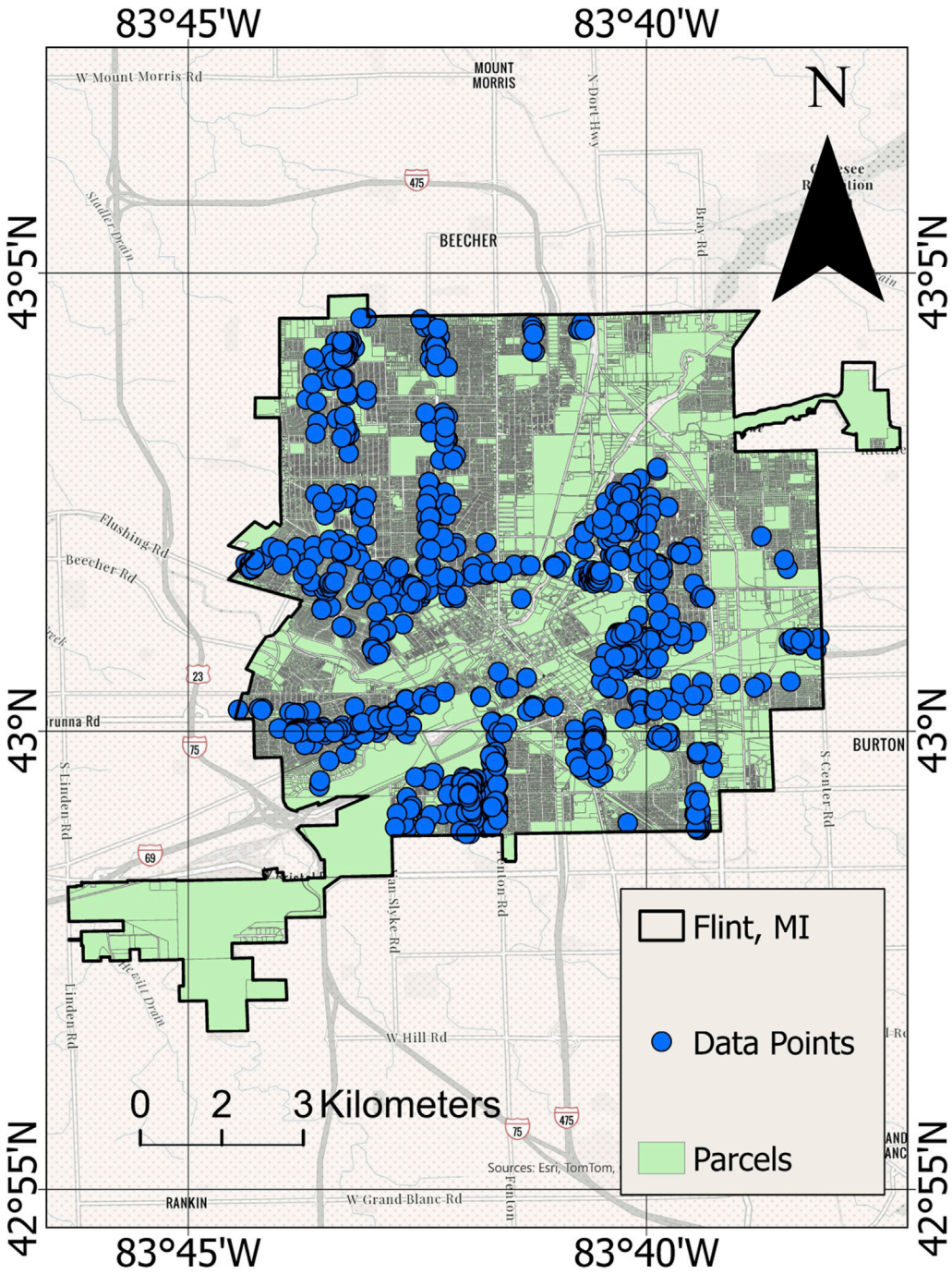


FIGURE 3
GIS map showing sampled houses in Flint for prediction studies.

home had consistent and verified data across all sources (Michigan Department of Environmental dataset, Zillow, and Arc GIS records). Samples that lacked corresponding information or exhibited more than 80% missing data were excluded. This stringent cleaning process ensured a quality dataset for modeling. Each lead sample was also matched to a specific parcel of land in Flint, and one-hot encoding and normalization techniques were applied to the final dataset, following best practices outlined in Kuhn and Johnson (2013).

Each dataset was categorized based on lead contamination levels according to the EPA standards Table 2. The lead levels are divided into four categories:

Proportions and Counts.

- Level 0: 59.5% of the homes (545 homes) are categorized as Level 0, indicating minimal lead contamination. Lead levels less than 5 ppb
- Level 0: 59.5% of the homes (545 homes) are categorized as Level 0, indicating minimal lead ad levels between 5 and 10 ppb

10.3389/fenve.2025.1488965

TABLE 1 Lead contamination range (ppb) and categorization.

Type	Lead value range (ppb)
Classifier 1	Less than 5 ppb
	5 ppb to less than 10 ppb
Classifier 2	Less than 5 ppb
	10 ppb to less than 15 ppb
Classifier 3	Less than 5 ppb
	15 ppb or greater

TABLE 2 Lead contamination range and implications.

Lead level (ppb)	Implications
0–5	None detectable
5–10	Minimal contamination, yet no safe level for children
10–15	Violates trigger level under the Revised Lead Rule; action required
>15	Exceeds EPA action level, indicating significant contamination and need for immediate action

- Level 0: 59.5% of the homes (545 homes) are categorized as Level 0, indicating minimal lead Level 2: 8.2% of the homes (75 homes) are in Level 2, with moderate lead contamination. Lead levels between 10 and 15 ppb
- Level 0: 59.5% of the homes (545 homes) are categorized as Level 0, indicating minimal lead Level 3: 15.6% of the homes (143 homes) are classified as Level 3, showing the highest levels of lead contamination. Lead levels 15 ppb or greater

The dataset is predominantly composed of homes with very low or no lead contamination (Level 0). However, a significant portion of the dataset (approximately 40.5%) consists of homes with varying degrees of lead contamination (Levels 1, 2, and 3). This stratification was crucial for ensuring the model could learn to identify both contaminated and homes at very low risk of contamination effectively. While minor seasonal variations in lead levels were observed, they do not significantly impact the overall assessment of contamination risk. The seasonal trends suggest that lead levels remain relatively stable throughout the year, with slight increases in warmer months. This reinforces the importance of long-term monitoring beyond seasonal patterns, as lead contamination is influenced by multiple factors beyond temperature fluctuations. Although this dataset does not represent the entire population of Flint homes, it provides a robust sample for testing and refining the model. The inclusion criteria ensured data quality and reliability, while the prioritization of contaminated homes allowed the model to focus on identifying risk patterns. This study lays the foundation for future work with larger, more representative datasets.

In summary, we utilize data from Flint, Michigan's lead contamination crisis to develop a machine-learning model for the detection of lead in drinking water.

2.4 Machine learning models

We employed GAT for our primary model, developed using libraries including PyTorch, PyTorch Geometric, NetworkX, and GeoPy (Veličković et al., 2018). For ensemble modeling, we used classical machine learning algorithms such as RandomForestClassifier, SVC, and XGBoost (Abernethy et al., 2016). The models' performance was assessed using metrics like the Receiver Operating Characteristic (ROC) curve, ROC, and the AUC score. Throughout this process, we adhered to ethical guidelines, particularly ensuring that data was not publicly accessible online, to safeguard our dataset's privacy and ethical integrity (Floridi and Taddeo, 2016).

2.5 Graph attention network

The Graph Attention Network (GAT), a specialized variant of Graph Neural Networks (GNNs), incorporates an attention mechanism to emphasize the influence of specific neighboring nodes. This approach was introduced in the study by Veličković et al. (2018). The attention mechanism, originally developed for sequence-based tasks, has been applied in areas such as machine translation (Bahdanau and Bengio, 2015; Vaswani et al., 2017). More recently, this concept has been adapted for graph-based applications, resulting in various models that integrate the attention operator into graph neural networks. The work by Veličković et al. (2018) represents an important step in extending attention-based methods to graph-structured data.

In this study, we utilize the GAT model to generate higher-level feature representations. The model applies self-attention mechanisms to an input graph composed of N nodes, each with F. Our implementation diverges from the original architecture by employing a single attention head, simplifying computations without significantly compromising performance (Veličković et al., 2018). The input to the GAT layer is a set of node features, represented as shown in Equations 1–4:

$$h = \{\vec{h}_1, \vec{h}_1, \ldots, \vec{h}_N,\} \tag{1}$$

Here $\vec{h}_i \in \mathbb{R}^F$. The layer produces a new set of node features, potentially with different cardinality $\mathbf{F}'$, denoted by $\mathbf{h}' = \{\mathbf{h}'_1, \mathbf{h}'_2, \ldots, \mathbf{h}_{\hat{N}}\}$, $\mathbf{h}'_i \in \mathbb{R}^{\hat{F}}$ as output. A shared linear transformation, parametrized by a weight matrix $\mathbf{W} \in \mathbb{R}^{\hat{F}} \times F$, is applied to every node. This transformation allows the network to express higher-level features. The self-attention mechanism computes attention coefficients $e_{ij} = a(W\vec{h}_i, W\vec{h}_j)$, reflecting the importance of node j's features to node i. The model includes masked attention, considering only nodes $j \in N_i$, where N_i refers to the first-order neighbors of i (including i). Attention coefficients are normalized across nodes using the softmax function:

$$\alpha_{ij} = \text{softmax}_j\left(e_{ij}\right) = \frac{\exp\left(e_{ij}\right)}{\sum_{k \in N_i} \exp\left(e_{ik}\right)}. \tag{2}$$

The attention mechanism a is modeled as a single-layer feedforward neural network, parametrized by a weight vector

$\vec{a} \in \mathbb{R}^{2F'}$, using a LeakyReLU nonlinearity with negative input slope $\alpha = 0.2$. The coefficients are computed as:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\vec{a}^T\left[W\vec{h}_i \| W\vec{h}_j\right]\right)\right)}{\sum_{k\in N_i}\exp\left(\text{LeakyReLU}\left(\vec{a}^T\left[W\vec{h}_i \| W\vec{h}_k\right]\right)\right)}. \quad (3)$$

Here $\cdot^T$ represents transposition and $\|$ is the concatenation operation. After the normalized attention coefficients are calculated, they are used to form a linear combination of the features that correspond to them, generating the final output features for every node:

$$\vec{h}_i' = \sigma\left(\sum_{j\in N_i}\alpha_{ij}W\vec{h}_j\right), \quad (4)$$

where σ is the activation function (e.g., ReLU or softmax).

2.6 Advantages and challenges of graph attention networks

GATs have emerged as an effective tool for handling graph-structured data, offering numerous benefits that are particularly relevant to tasks requiring nuanced relational modeling. At the heart of GATs lies their attention mechanisms, which selectively prioritize important nodes and edges within a graph. This selective focus enhances computational efficiency through parallel processing while simultaneously boosting the model's ability to extract and emphasize critical relationships. Moreover, the inductive learning capabilities of GATs, enabled by their shared attention mechanisms, extend their versatility to a wide range of scenario. The interpretability of GATs is another significant advantage; by providing insights into the decision making process through learned attention weights. Additionally, their inherent flexibility in accommodating dynamic structures has made them good choice for graph-based data analysis (Zhou et al., 2020). Nevertheless, despite these strengths, GATs are not without challenges. They are often hindered by scalability issues, computational intensity, and a susceptibility to overfitting (Vrahatis et al., 2024). Recognizing these limitations, we adopted several measures to address these concerns and optimize the performance of GATs in our study. To manage computational demands, we simplified the architecture by employing a single attention head instead of multiple, thereby streamlining computations while maintaining effective performance. Furthermore, the stability of the model was enhanced through the use of LeakyReLU activations, which mitigated the risk of gradient vanishing during softmax normalization and ensured reliable convergence throughout the training process. In addition, to address the challenges posed by graph density, we optimized graph construction by connecting nodes based on meaningful geodesic distance thresholds. This approach not only reduced the overall sparsity of the graph but also preserved essential spatial relationships, thereby improving computational efficiency without compromising the quality of the relational data captured. By implementing these targeted strategies, we tailored the GAT model to align with the specific requirements of our dataset and problem context. Ultimately, while GATs excel in capturing the complexities of graph-structured data, these adjustments underscore the importance of acknowledging their limitations. Our approach strikes a balance between leveraging their inherent strengths and mitigating their constraints, ensuring that GATs remain a practical and effective choice for our study without overstating their universal applicability.

2.7 Evaluation metrics

In the field of machine learning, evaluation metrics play a pivotal role in assessing the performance and reliability of the models. The Receiver Operating Characteristic (ROC) curve serves as a critical tool for evaluating classification performance across various decision thresholds of the model (Fawcett, 2006). It represents the trade-off between the True Positive Rate (TPR), or sensitivity, and the False Positive Rate (FPR), defined as shown in Equation 5:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \quad (5)$$

where TP, FN, FP, and TN correspond to true positives, false negatives, false positives, and true negatives, respectively. The ROC curve is instrumental in evaluating a model's ability to balance sensitivity and specificity. Additionally, the Area Under the Curve (AUC) quantifies the model's overall discriminatory power and is calculated using Equation 6:

$$\text{AUC} = \int_0^1 \text{TPR}\left(\text{FPR}\right) d\left(\text{FPR}\right), \quad (6)$$

where an AUC of 0.5 indicates random chance, while a value closer to 1.0 signifies exemplary classification performance.

2.7.1 Cross-validation and stability

To ensure the robustness of the proposed model, we employed 10 independent runs with cross-validation, calculating the mean accuracy and standard deviation as measures of prediction consistency. These metrics were computed as shown in Equation 7:

$$\text{Mean Accuracy} = \frac{1}{n}\sum_{i=1}^{n}\text{Accuracy}_i,$$

$$\text{Standard Deviation} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\text{Accuracy}_i - \text{Mean Accuracy}\right)^2}, \quad (7)$$

where $n = 10$ denotes the number of runs. This approach highlighted the model's stability and minimized concerns regarding overfitting or variability across training sessions. By including standard deviation as a performance measure, we ensured reliable predictions under varying conditions.

2.7.1.1 Probability distribution

A softmax function was applied to the output layer of the model to normalize raw scores into interpretable probabilities as defined in Equation 8:

$$P\left(y = c|x\right) = \frac{\exp\left(z_c\right)}{\sum_{k=1}^{C}\exp\left(z_k\right)}, \quad (8)$$

where z_c represents the logit score for class c, and C denotes the total number of classes.

2.7.1.2 Feature normalization

Input features were scaled to comparable ranges, promoting numerical stability and preventing biases from dominating training. This scaling was achieved using Equation 9:

$$x_{\text{norm}} = \frac{x - \mu}{\sigma}, \tag{9}$$

where μ is the mean, and σ is the standard deviation of the feature.

2.8 Addressing overfitting and computational considerations

In our model we mitigated Overfitting through the integration of dropout and weight decay. Dropout was applied at a rate of $p = 0.6$, reducing the risk of overfitting by randomly excluding neurons during training. Additionally, weight decay ($\lambda = 0.001$) provided L2 regularization, further enhancing the model's generalization capabilities.

Numerical stability, particularly in the self-attention mechanisms of the Graph Attention Network (GAT), was maintained through the use of LeakyReLU activations, as shown in Equation 10:

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x \geq 0, \\ \alpha x & \text{if } x < 0, \end{cases} \tag{10}$$

with $\alpha = 0.2$. This ensured non-zero gradients, preserving learning dynamics even during softmax normalization. Careful parameter initialization further stabilized training, particularly in the early stages, reducing the risk of vanishing or exploding gradients.

2.9 Baseline ensemble approach

Ensemble learning is a principled approach in machine learning that combines predictions from multiple models to achieve improved predictive accuracy. This model implementation closely follows the approach of Abernethy et al. (2016), as we aim to compare their XGBoost approach to our GAT-based methodology. Figure 4 shows the flow chart of the XGBoost method. This section delves into the techniques and mathematical formulations used in the ensemble learning methodology, particularly focusing on the stacking technique, models employed, evaluation metrics, and calibration measures. The ensemble learning method used constructs a predictive model by aggregating predictions from a collection of individual models. The combined prediction can be expressed mathematically in Equation 11:

$$f(x) = \sum_{i=1}^{M} w_i \cdot f_i(x) \tag{11}$$

Here, $f(x)$ denotes the ensemble prediction, M is the number of models, w_i are the weights, and $f_i(x)$ are the individual model predictions. The given model adopts stacking, a popular ensemble technique that uses predictions from various models (first layer) which includes (Chen and Guestrin, 2016), random forest (Breiman, 2001), extremely randomized trees(Geurts et al., 2006), logistic regression (Fisher, 1936), nearest neighbor (Cover and Hart, 1967), and linear discriminant analysis (LDA) (Guisan et al., 2002) as input and a second layer of a single XGBoost classifier for ensembling.

The first layer trains multiple classifiers on the dataset. Each model, $f_i(x)$, produces a prediction, and these predictions are then stacked together as shown in Equations 12–13:

$$P = [p_1, p_2, \ldots, p_M] \tag{12}$$

here p_i is the prediction of the i-th model.

The second layer is responsible for training the final model (e.g., XGBoost) on the stacked predictions P to form the final prediction:

$$f(x) = g(P) \tag{13}$$

here g is the second-layer model.

The ensemble's performance is assessed using the ROC-AUC and a confusion matrix. It measures the area under the Receiver Operating Characteristic curve, representing the model's ability to discriminate between positive and negative classes.

3 Results and discussion

3.1 Overview of results

This study used a Graph Attention Network (GAT) to predict lead contamination levels in Flint, Michigan, focusing on contamination thresholds between 5 and 15 ppb. Our model performed well, with an Area Under the Curve (AUC) of 0.93, precision of 71.55%, and recall of 93.77%. In contrast, the XGBoost implementation had an AUC of 0.66, indicating its limitations in capturing complex spatial and relational data patterns.

These results highlight the effectiveness of GAT in modeling lead contamination, especially in incorporating spatial features like parcel adjacency and housing proximity. Figure 5 illustrates the ROC curve for our model tested on the test dataset. The model, incorporates data from both the parcel layer and the Zillow data, achieving an impressive Area Under the Curve (AUC) of 0.93. This AUC value, combined with its proximity to the top-left corner of the graph, demonstrates the model's efficiency in classification, indicated by a high TPR and a low FPR.

3.2 Comparative analysis with existing literature

We conducted a comparative analysis with key studies in the field to contextualize our findings within broader lead contamination research. This comparison underscores the progress achieved by our GAT model and its strengths in using spatial and relational data for better predictive performance. Abernethy et al. (2016) used an ensemble method to identify lead service lines in Flint, achieving an accuracy of 0.677. Although their approach demonstrated the potential of ensemble methods helping to identify lead, our GAT model's AUC of 0.93 indicates a significant improvement. This suggests the importance of incorporating spatial

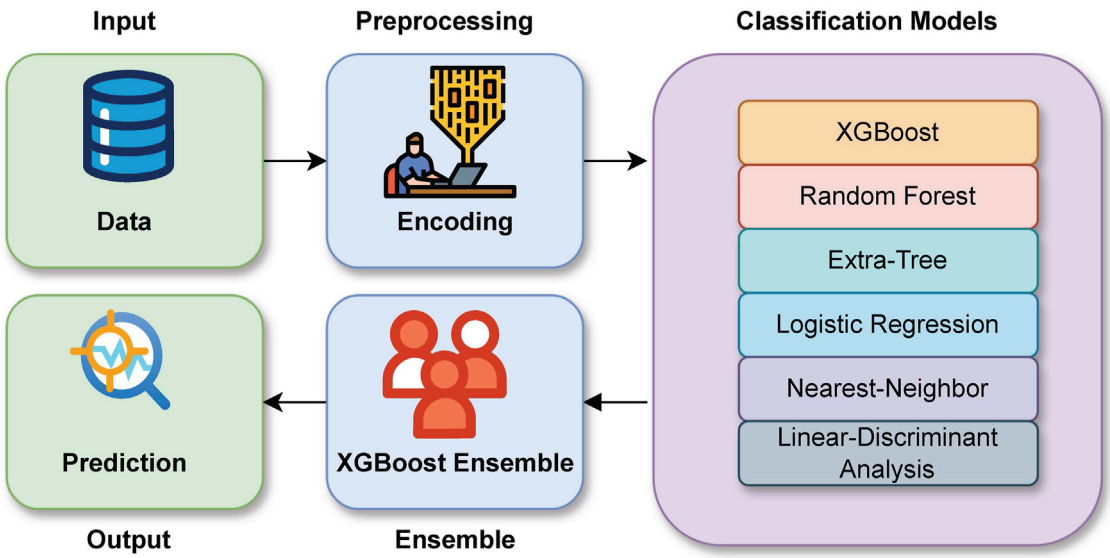


FIGURE 4
Flowchart of Ensemble-based Method used in the XGBoost algorithm Method WebHarvy (2022).

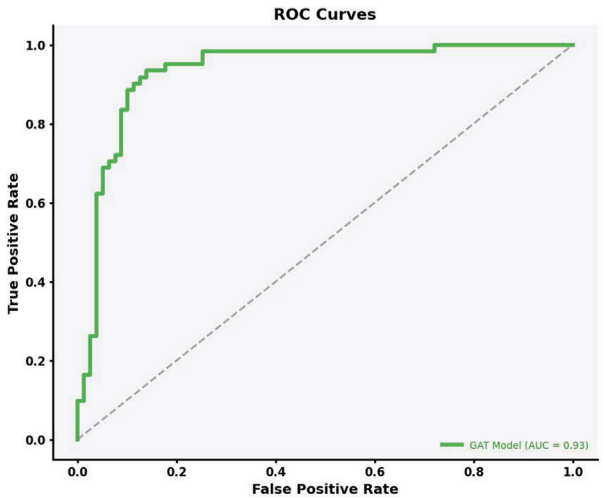


FIGURE 5
True positive vs. False positive rates ROC curve for GAT Model.

and relational features, which are key strengths of GAT. Similarly, Goovaerts (2019) used a multivariate geostatistical approach (cokriging) to predict water lead levels, achieving an AUC of 0.76. Cokriging effectively captures spatial variability by integrating multiple data sources. Despite this, our GAT model achieved a higher AUC. While Goovaerts' work offers very valuable insights into data integration, our results show that graph based modeling might provide a robust framework for predicting lead contamination. Mulhern et al. (2023) applied Bayesian Networks (BN) to predict lead risk, achieving an AUC of 0.74. Their method was particularly good at identifying high-risk facilities with clustered contamination. However, our GAT model showed higher AUC and recall, reflecting its ability to identify contaminated parcels more effectively. GAT's capacity to model complex spatial relationships gives it broader applicability, particularly in urban areas where contamination is influenced by neighboring properties. In the study by Early Warning Systems (Khaksar Fasaee et al., 2022), Bayesian classifiers and Ensemble Decision Trees (EDT) were used to predict lead contamination in private water systems, achieving an AUC of 0.77 and recall of 75%. Although these models were effective, especially when incorporating household-level features, our GAT model achieved a higher AUC and recall. Overall, these comparisons show that our GAT model not only achieves high performance but also addresses some of the limitations in previous methods. By effectively integrating spatial and relational data, our approach improves predictive accuracy of the previous studies. This methodology offers a tool for lead contamination mitigation, ensuring resources are efficiently allocated to at risk households. The information presented in this study contribute to the ongoing research on lead contamination prediction and highlight the potential of using graph based methods in environmental health.

In Figure 6, the graph shows how the accuracy of the model changes as the distance threshold for connecting houses in the graph increases. Here, the threshold is the maximum distance between houses, in miles, that defines whether they are considered connected in the Graph Attention Network (GAT). At a 0.1-mile threshold, houses are connected only to their closest neighbors, and this setting yields the highest model accuracy (ACC), slightly above 80%. As the distance threshold increases to 0.3 miles, the model accuracy gradually decreases to just above 77%. This indicates that the model benefits from focusing on more local relationships. The trade off seen in the graph highlights the importance of selecting an appropriate threshold to find the right balance between connectivity and prediction quality. The confusion matrix results for the different thresholds are summarized in Table 3.

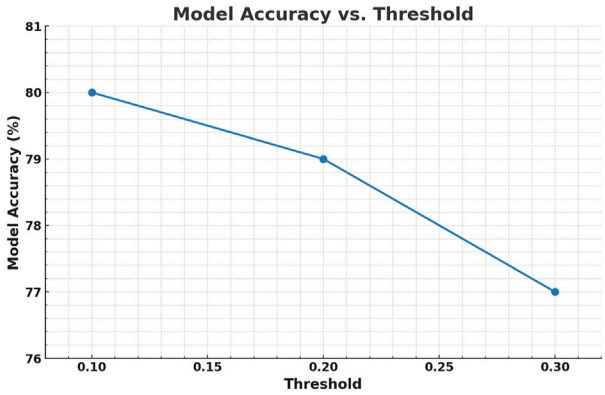


FIGURE 6
GAT model accuracy across different threshold.

TABLE 3 Performance metrics based on distance threshold GAT model.

Distance Threshold	Performance metrics (%)			
	TP	FP	FN	TN
0.1	96.15	3.28	28.57	71.43
0.2	96.72	3.28	24.05	75.95
0.3	95.08	5.45	26.58	73.42

3.3 Ensemble learning with XGBoost

In addition to our evaluation of the GAT model, we compared the performance with an ensemble learning approach using the XGBoost method, a gradient-boosted decision tree algorithm, which was previously used by Abernethy et al. (2016). Ensemble models like XGBoost combine multiple individual models to achieve better predictive performance compared to individual models. XGBoost, in particular, has gained recognition for its ability to handle large and unbalanced datasets effectively. Abernethy et al. (2016) employed a model to identify lead service lines in Flint, reporting an accuracy of 67.7% and a log loss of 0.054. In our study, the XGBoost model achieved an accuracy of 70.0%, a precision of 65.57%, and a recall of 66.67%. While these results were comparable to those reported by Abernethy and Yang, the GAT model performed better across all key metrics. Specifically, the GAT model achieved an AUC of 0.93, a precision of 71.55%, and a recall of 93.77% Table 4. This demonstrates that the GAT model was better in identifying and correctly labeling homes that might have lead contamination. The confusion matrix in Figure 7 further illustrates the classification performance of the GAT model, showing the distribution of correct and misclassified predictions for both models. The main reason for this difference is that GAT can better leverage relationships between neighboring houses, which appears to be crucial for predicting lead contamination accurately. XGBoost, in contrast, focuses more on the volume and diversity of data without directly incorporating spatial relationships, which could explain why the GAT model proved to be effective.

Although the GAT model had higher accuracies than XGBoost, XGBoost model still has strengths, such as its resistance to overfitting and scalability, which is a common advantage of gradient-boosted models. This was particularly useful in our case, where the dataset contained a lot of features. Moreover, the XGBoost model offered insights into which features were important, such as Housing Age and Property Address Street, which were also identified as significant by the GAT model. While GAT uses the spatial relationships between houses, XGBoost did well with the volume and variety of data. This difference explains their varying performance. In situations where spatial relationships are not the most critical aspect, XGBoost might still be a good alternative or complement to GAT. Future research could also explore combining both models to see if that improves its accuracy.

3.4 Performance comparison of GAT and XGBoost across different contamination thresholds

To show the performance difference between GAT and XGBoost, we turned the classification into a binary problem, using different lead contamination thresholds (5 ppb, 10 ppb, and 15 ppb). The GAT model consistently had higher accuracies compared to the XGBoost across all thresholds.

3.5 Sensitivity analysis

To further evaluate (GAT) model, we conducted sensitivity analyses, including an assessment of the ROC curve across contamination thresholds and an examination of different spatial distance thresholds.

3.6 ROC curve analysis

The ROC curve was used to assess how well the GAT model balances sensitivity and specificity across different thresholds. The model achieved an AUC of 0.93, showing its ability to distinguish between contaminated and non-contaminated parcels. This performance was better than that of XGBoost

TABLE 4 Performance Metrics for Lead Contamination Prediction Models.

Model and data source	Accuracy (%)	Precision (%)	Recall (%)
XGBoost Ensemble (Parcel Data)	70.0	65.57	66.67
GAT (Parcel and Zillow Data)	80.57	71.55	93.77

The table shows accuracy, precision, and recall for each model. Accuracy is the percentage of correct predictions, precision indicates the correctness of positive predictions, and recall measures how well the model identifies all positive cases.

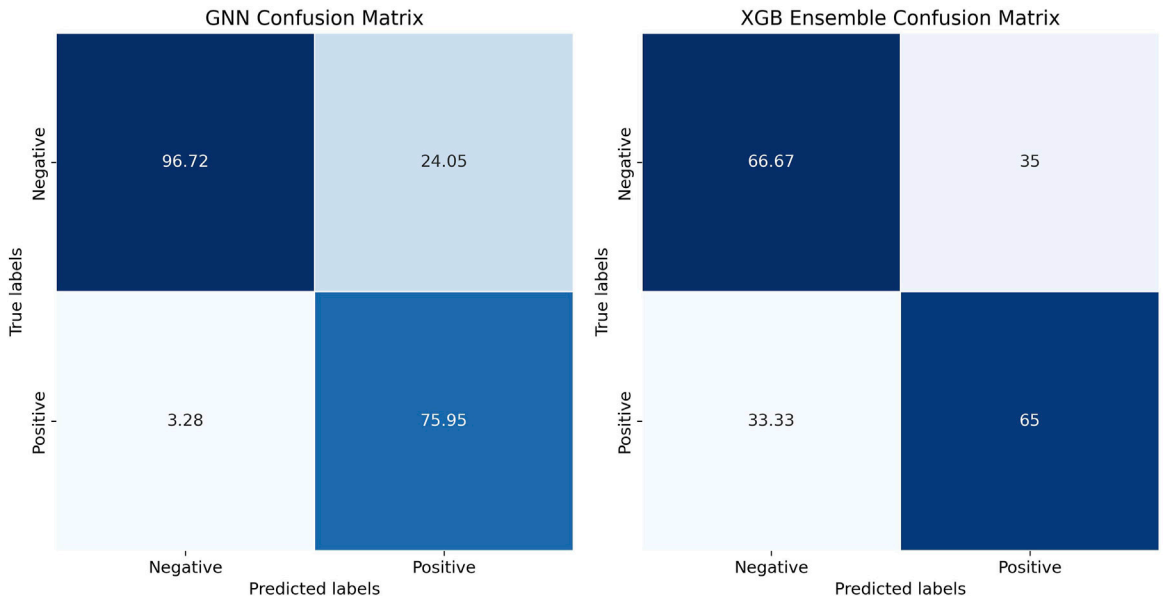


FIGURE 7
Confusion matrix comparison for model performance.

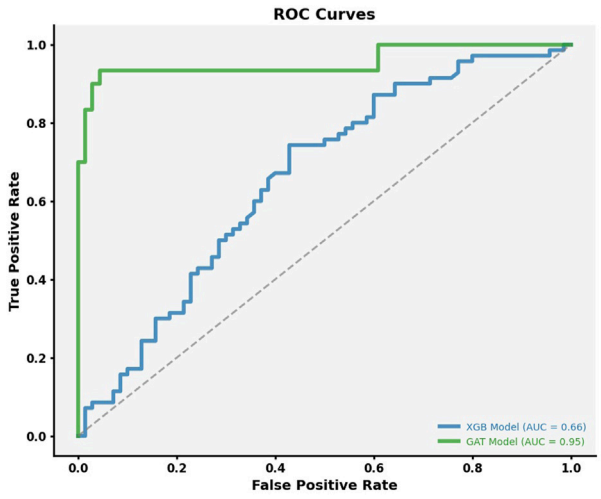


FIGURE 8
ROC Curve for Classifier 3 (15 ppb): Evaluating Model Performance at the EPA Action Level. This curve demonstrates the model's ability to distinguish between houses with lead contamination levels above and below the EPA action threshold of 15 ppb.

(AUC: 0.66) and Bayesian Networks (AUC: 0.74), highlighting GAT's effectiveness in handling imbalanced datasets and capturing spatial relationships. Figure 5 shows the ROC curve for the GAT model across multiple thresholds, while Figure 8 focuses specifically on classifier performance at the EPA action level of 15 ppb. The model's proximity to the left-top corner in these curves indicates strong predictive accuracy. In practical terms, this means that the GAT model can identify contaminated households effectively while reducing misclassification, which is important for managing limited resources in lead mitigation efforts.

3.7 Spatial distance threshold analysis

To complement the ROC curve analysis, we also evaluated how different spatial distance thresholds affected recall and precision. The GAT model was tested with thresholds ranging from 0.1 to 0.3 miles, which represented different levels of connectivity within the graph. At a 0.1-mile threshold, the model achieved the highest recall (93.77%), ensuring that most contaminated parcels were identified. Increasing the threshold to 0.3 miles slightly reduced recall but improved precision. This trade off highlights the spatial sensitivity of the model and the importance of selecting appropriate thresholds for specific public health goals. While spatial threshold analysis offers useful insights for optimizing the graph structure, the ROC curve analysis remains the primary metric for evaluating overall model performance. The ROC curve effectively summarizes the trade off between sensitivity and specificity across different decision thresholds.

3.8 Practical applications and implications

In public health contexts, such as the Flint water crisis, accurate predictive models are crucial for safety. A model that can predict lead levels in residential water helps guide interventions, ensuring that resources are directed to the areas that need them most. The aim is not just to achieve good accuracy but to reduce health risks for residents. Traditional methods often struggle to capture the spatial dependencies in lead contamination data. The GAT model effectively uses these spatial relationships, which allows for a better understanding of how contamination spreads. Future work could explore hybrid models that combine GAT with

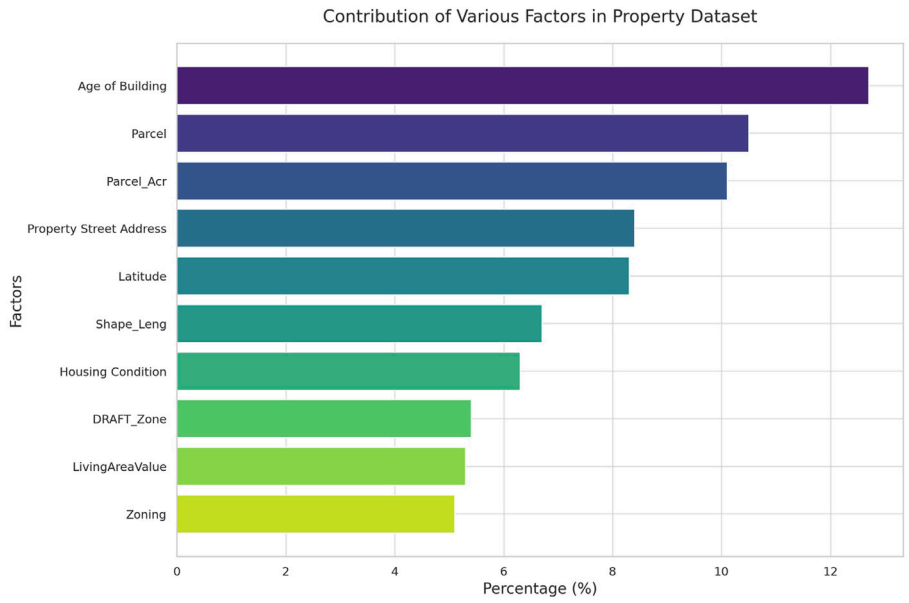


FIGURE 9
Top-predicting features.

statistical approaches, like cokriging, for more predictive accuracy. Additionally, the model's predictions can be used to help strategize local interventions, ensuring efficient use of resources.

3.9 Relative variable influence

In our research utilizing the GAT framework, we aimed to pinpoint key node features crucial for predicting homes at risk for lead contamination in water. It is important to note that the identified risk factors do not directly imply causation of lead contamination but are useful in differentiating parcels with a higher likelihood of unsafe lead levels. By applying the GAT framework, which is effective in handling relational data, we identified features such as the Age of Building, Property Address, Zoning, and Parcel as significant predictors. For instance, older buildings may have outdated plumbing that increases lead risk. These findings could lead to more targeted inspections and informed policy-making aimed at reducing lead contamination.

To explain the importance of each feature, we used GraphLIME, a local interpretable model-agnostic explanation method tailored to Graph Neural Networks. This method approximates the complex GNN model with a simpler, interpretable model for a specific node's neighborhood. By perturbing the node features and observing the changes in the model's output, we determined the importance score of each feature:

$$I(f_i) = P(F) - P(F \backslash f_i)$$

Where.

- $I(f_i)$ represents the importance score of feature f_i.
- F is the complete set of features.
- $F \backslash f_i$ denotes the feature set excluding feature f_i.

Figure 9 shows the top important features as determined by our model's output.

Interestingly, our model revealed that some features traditionally considered important in lead contamination studies, such as Longitude Elevation, were less influential. This suggests a more complex role for geographic factors in lead contamination risk. Features like DraftZone and Housing Condition (HCond 2012) were more influential, emphasizing the importance of housing-related factors in lead risk. Even features that might seem less critical, such as Living Area Value and Parcel Acres, provide valuable context for understanding potential lead contamination sources. These insights can help guide targeted and effective interventions, shaping policies to prioritize resources for mitigating lead contamination.

3.10 Practical applications and broader context

The accuracy of predictive models in public health, particularly in scenarios like Flint's lead crisis, has real implications for community safety. Correctly predicting lead levels in residential water allows authorities to plan interventions and allocate resources efficiently, focusing on the most vulnerable areas first.

The ROC curves in Figure 10 provide additional insight into the model's classification performance at these thresholds. Our model consistently performed better than traditional methods in identifying contaminated homes, particularly across different lead thresholds (>5, >10, >15 ppb), as shown in Figure 11. The use of GAT captures spatial dependencies and relationships effectively, something that traditional methods, including Flint's ensemble models, struggled to do. The attention mechanism of GAT allows it to weigh neighboring nodes differently, which leads to a better understanding of how lead contamination spreads or clusters.

The success of integrating alternative data sources, such as Zillow information, highlights the potential for unconventional but relevant data to provide additional context in a crisis. It also

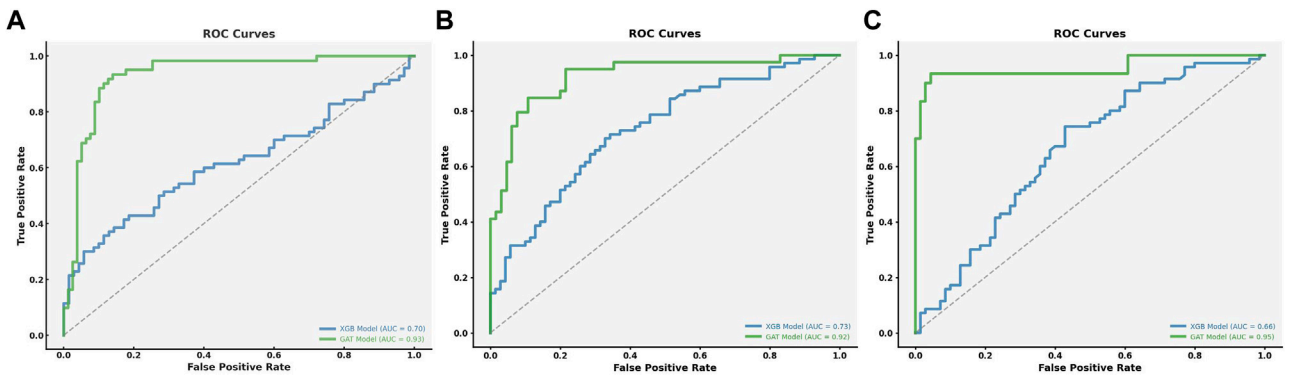


FIGURE 10
ROC curves of true positive rates vs. false positive rates for the models with different thresholds. The first figure represents a threshold of >5 ppb, the second figure represents a threshold of >10 ppb, and the third figure represents a threshold of >15 ppb. These figures illustrate the model's ability to distinguish between contaminated and non-contaminated houses at different levels of lead concentration.

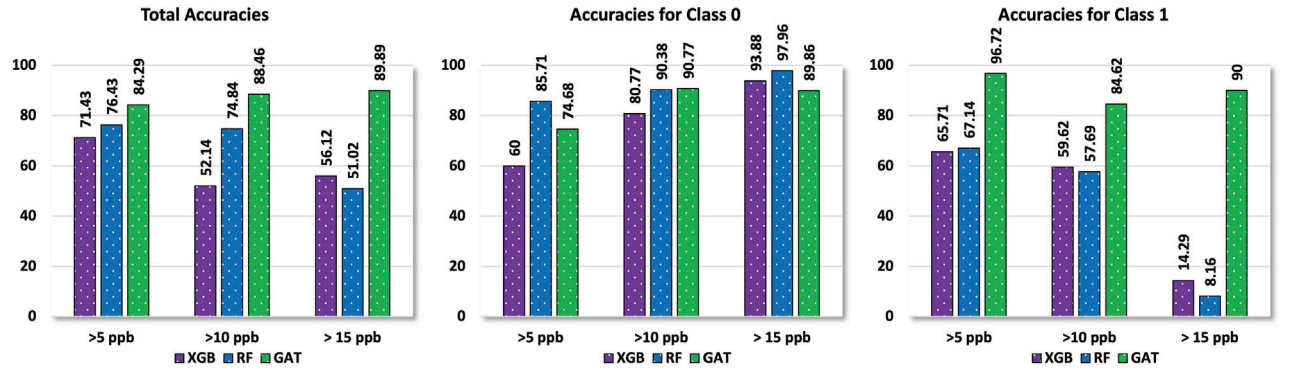


FIGURE 11
Our algorithm's binary classification performance comparisons for different thresholds with different baseline algorithms. The left figure shows overall performance, while the middle and right figures show performance for each class (contaminated vs. non-contaminated) for thresholds (>5, >10, >15 ppb).

shows the importance of collaboration between public entities and private data holders to enrich available data and improve model accuracy. Future research could explore how other data sources might be leveraged in similar contexts, whether for pollution monitoring, understanding disease spread, or urban planning.

3.11 Applicability beyond the USA

Although this model was developed using U.S. data, applying it elsewhere requires adjusting for local regulations, infrastructure, and data availability. Different regions may have stricter or more lenient lead thresholds, unique sources of contamination (e.g., industrial pollution or agricultural runoff), and varying monitoring practices. Retraining or recalibrating the model with local data is often necessary to capture these differences accurately. In areas lacking centralized water quality records, community-driven sampling or municipal reports might be required. The lessons from Flint demonstrate how data-driven approaches can support public health interventions, reinforcing the importance of predictive models that integrate machine learning with real-world applications.

4 Conclusion

The Flint water crisis is a clear reminder of the serious consequences that happen when environmental and public health problems collide. This study shows that Graph Attention Networks (GAT) can be a valuable tool to help predict lead contamination more accurately, thanks to its ability to capture complex spatial relationships. However, it's important to note that the GAT model is just one part of the solution it complements other efforts and supports better decision making.

One of the key findings of this study is the identification of features that are strong predictors of high lead levels in homes, even without knowing the exact composition of Lead Service Lines (LSLs). This is crucial information that can help policymakers and community members target high-risk areas, allocate resources more efficiently, and take action to protect public health.

Technical Implications: The GAT model has shown strong potential in improving how we assess water quality risks, helping us identify high-risk areas more effectively. Municipalities could use similar models to focus their efforts on the neighborhoods most vulnerable to lead contamination. Future monitoring systems

should look to integrate such graph-based tools to get ahead of issues before they escalate.

Policy Implications: The findings suggest that incorporating predictive models like GAT can help public health authorities make better use of limited resources. By pinpointing areas most likely to face lead contamination, actions can be prioritized to prevent health risks before they become widespread. Additionally, this research highlights the value of using diverse data sources—like real estate information—to enrich monitoring efforts. Collaborations between public and private entities can make datasets more robust and improve the ability to address contamination proactively.

Limitations: While the GAT model offers many benefits, it's not without its limitations. The model relies heavily on the quality and availability of data as well as scalability. Users of this model should be mindful of these constraints and recognize that the model should be seen as a guiding tool rather than a definitive solution. Continuous data collection, validation, and refinement of the model are needed to improve accuracy and reduce uncertainties.

The findings from this research are not just relevant to Flint. The approach used in this study can be applied in other areas, like urban planning, pollution monitoring, and public health research. By bringing together public health agencies, local governments, and private data providers, we can create more effective datasets that ultimately lead to safer communities.

In conclusion, addressing public health issues like lead contamination requires a combination of advanced technology and practical policy changes. By using tools like GAT alongside careful policymaking and resource allocation, we can tackle challenges like the Flint water crisis more effectively. The future focus should be on refining these models, improving data quality, and extending these approaches to other communities to better manage lead contamination risks on a broader scale.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.michigan.gov/flintwater/water-infrastructure-projects/monitoring/wqa-flint-dist/city-of-flint-distribution-system-monitoring-data-expanded-data-set.

Author contributions

RA: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. YB: Conceptualization, Data curation, Investigation, Visualization, Writing–review and editing. CS: Visualization, Writing–review and editing. MP: Data curation, Writing–review and editing. PK: Formal Analysis, Funding acquisition, Project administration, Resources, Writing–review and editing. BL: Funding acquisition, Writing–review and editing, Project administration. MA: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by a grant from the National Science Foundation (NSF) under the Smart and Connected Community (SCC) program. The grant, approximately valued at $2.5 million USD, is titled "SCC-IRG Track 1: Community Based Approach to Address Contaminants in Drinking Water using Smart Cloud-Connected Electrochemical Sensors," with the award number 2230180.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenve.2025.1488965/full#supplementary-material

References

Abernethy, J., Anderson, C., Dai, C., Farahi, A., Nguyen, L., Rauh, A., et al. (2016). Flint water crisis: data-driven risk assessment via residential water testing. doi:10.48550/arXiv.1610.00580

Agency, U. E. P. (2023). Lead regulations.

Aldhyani, T. H. H., Al-Yaari, M., Alkahtani, H., and Maashi, M. (2020). Water quality prediction using artificial intelligence algorithms. *Appl. Bionics Biomechanics* 2020, 1–12. doi:10.1155/2020/6659314

Bae, Y., Pasteris, J. D., and Giammar, D. E. (2020a). The ability of phosphate to prevent lead release from pipe scale when switching from free chlorine to monochloramine. *Environ. Sci. and Technol.* 54, 879–888. doi:10.1021/acs.est.9b06019

Bae, Y., Pasteris, J. D., and Giammar, D. E. (2020b). Impact of iron-rich scale in service lines on lead release to water. *AWWA Water Sci.* 2, e1188. doi:10.1002/aws2.1188

Bae, Y., Pasteris, J. D., and Giammar, D. E. (2020c). Impact of orthophosphate on lead release from pipe scale in high ph, low alkalinity water. *Water Res.* 177, 115764. doi:10.1016/j.watres.2020.115764

Bahdanau, C. K. D., and Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. doi:10.48550/arXiv.1409.0473

Bai, T., and Tahmasebi, P. (2023). Graph neural network for groundwater level forecasting. *J. Hydrology* 616, 128792. doi:10.1016/j.jhydrol.2022.128792

Blackhurst, M. (2021). Identifying lead service lines with field tap water sampling. *ACS ES&T Water* 1, 1983–1991. doi:10.1021/acsestwater.1c00227

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi:10.1023/A:1010933404324

Chen, T., and Guestrin, C. (2016). Xgboost: a scalable tree boosting system. *CoRR*. doi:10.48550/arXiv.1603.02754

Chojnacki, A., Dai, C., Farahi, A., Shi, G., Webb, J., Zhang, D. T., et al. (2017). "A data science approach to understanding residential water contamination in flint," in Proceedings of the 23rd ACM SIGKDD international Conference on knowledge Discovery and data mining *(ACM)*. doi:10.1145/3097983.3098078

Cornwell, D. A., Brown, R. A., and Via, S. H. (2016). National survey of lead service line occurrence. *J. AWWA* 108, E182–E191. doi:10.5942/jawwa.2016.108.0086

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27. doi:10.1109/tit.1967.1053964

Devine, C., Mello, K., DeSantis, M., Schock, M., Tully, J., and Edwards, M. (2024). Calcium phosphate precipitation as an unintended consequence of phosphate dosing to high-ph water. *Environ. Eng. Sci.* 41, 171–179. doi:10.1089/ees.2023.0190

Doré, E., Lytle, D. A., Wasserstrom, L., Swertfeger, J., and Triantafyllidou, S. (2020). Field analyzers for lead quantification in drinking water samples. *Crit. Rev. Environ. Sci. Technol.* 51, 2357–2388. doi:10.1080/10643389.2020.1782654

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognit. Lett.* 27, 861–874. doi:10.1016/j.patrec.2005.10.010

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7, 179–188. doi:10.1111/j.1469-1809.1936.tb02137.x

Floridi, L., and Taddeo, M. (2016). What is data ethics? *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 374, 20160360. doi:10.1098/rsta.2016.0360

García-Timermans, C., Malfroot, B., Dierendonck, C., Mol, Z., Pluym, T., Waegenaar, F., et al. (2023). Pilot-scale drinking water distribution system to study water quality changes during transport. *npj Clean. Water* 6, 52. doi:10.1038/s41545-023-00264-8

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Mach. Learn.* 63, 3–42. doi:10.1007/s10994-006-6226-1

Goovaerts, P. (2019). Geostatistical prediction of water lead levels in flint, Michigan: a multivariate approach. *Sci. Total Environ.* 647, 1294–1304. doi:10.1016/j.scitotenv.2018.07.459

Guisan, A., Edwards, T. C., Jr, and Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* 157, 89–100. doi:10.1016/s0304-3800(02)00204-1

Gurewitsch, R. (2019). Pb-predict: using machine learning to locate lead plumbing in a large public water system

Hajiseyedjavadi, S., Blackhurst, M., and Karimi, H. A. (2020). A machine learning approach to identify houses with high lead tap water concentrations. *Proc. AAAI Conf. Artif. Intell.* 34, 13300–13305. doi:10.1609/aaai.v34i08.7040

Hajiseyedjavadi, S., Karimi, H. A., and Blackhurst, M. (2022). Predicting lead water service lateral locations: geospatial data science in support of municipal programming. *Socio-Economic Plan. Sci.* 82, 101277. doi:10.1016/j.seps.2022.101277

Han, H., Pan, D., Li, Y., Wang, J., and Wang, C. (2020). Stripping voltammetric determination of lead in coastal waters with a functional micro-needle electrode. *Front. Mar. Sci.* 7. doi:10.3389/fmars.2020.00196

Hensley, K., Bosscher, V., Triantafyllidou, S., and Lytle, D. A. (2021). Lead service line identification: a review of strategies and approaches. *AWWA Water Sci.* 3, 1–19. doi:10.1002/aws2.1226

Khaksar Fasaee, M. A., Pesantez, J., Pieper, K. J., Ling, E., Benham, B., Edwards, M., et al. (2022). Developing early warning systems to predict water lead levels in tap water for private systems. *Water Res.* 221, 118787. doi:10.1016/j.watres.2022.118787

Kuhn, M., and Johnson, K. (2013). *Applied predictive modeling*. Springer. doi:10.1007/978-1-4614-6849-3

Li, G., Bae, Y., Mishrra, A., Shi, B., and Giammar, D. E. (2020). Effect of aluminum on lead release to drinking water from scales of corrosion products. *Environ. Sci. and Technol.* 54, 6142–6151. doi:10.1021/acs.est.0c00738

Lin, Y., Qiao, J., Bi, J., Yuan, H., Gao, H., and Zhou, M. (2022). "Hybrid water quality prediction with graph attention and spatio-temporal fusion," in *2022 IEEE international conference on systems, man, and cybernetics (SMC)*, 1419–1424. doi:10.1109/SMC53654.2022.9945293

Lobo, G., Laraway, J., and Gadgil, A. (2022). Identifying schools at high-risk for elevated lead in drinking water using only publicly available data. *Sci. Total Environ.* 803, 150046. doi:10.1016/j.scitotenv.2021.150046

Lopez, K. G., Xiao, J., Crockett, C., Lytle, C., Grubbs, H., and Edwards, M. (2024). Zinc orthophosphate can reduce nitrate-induced corrosion of lead solder. *ACS ES&T Water* 4, 3153–3162. doi:10.1021/acsestwater.3c00786

Martin, S., and Griswold, W. (2009). Human health effects of heavy metals. *Environ. Sci. Technol. Briefs Citizens*, 1–6. doi:10.4236/jep.2017.811077

Michigan.gov (2023). Michigan lead safe home program.

Mulhern, R. E., Kondash, A., Norman, E., Johnson, J., Levine, K., McWilliams, A., et al. (2023). Improved decision making for water lead testing in u.s. child care facilities using machine-learned bayesian networks. *Environ. Sci. and Technol.* 57, 17959–17970. doi:10.1021/acs.est.2c07477

News21 (2023). Millions consumed potentially unsafe water in the last 10 years.

Ni, R., Chu, X., Liu, R., Shan, J., Tian, Y., and Zhao, W. (2024). Chromium immobilization and release by pipe scales in drinking water distribution systems: the impact of anions. *Sci. Total Environ.* 906, 167600. doi:10.1016/j.scitotenv.2023.167600

Pan, W., Pan, C., Bae, Y., and Giammar, D. (2019). Role of manganese in accelerating the oxidation of pb(ii) carbonate solids to pb(iv) oxide at drinking water conditions. *Environ. Sci. and Technol.* 53, 6699–6707. doi:10.1021/acs.est.8b07356

Pasteris, J., Bae, Y., Giammar, D., Dybing, S., Yoder, C., Zhao, J., et al. (2021). Worth a closer look: Raman spectra of lead-pipe scale. *Minerals* 11, 1047. doi:10.3390/min11101047

Potash, B., Brew, J., Loewi, A., Majumdar, S., Reece, A., Walsh, J., et al. (2015). "Predictive modeling for public health," in *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. doi:10.1145/2783258.2788629

Sandvig, A., Kwan, P., Kirrmeyer, G., Maynard, B., Mast, D., Trussell, R., et al. (2008). *Contribution of service line and plumbing fixtures to lead and copper rule compliance issues*. Denver, CO: Water Research Foundation.

Sarr, S. O. E. N., and Diallo, A. (2018). "Factextract: automatic collection and aggregation of articles and journalistic factual claims from online newspaper," in *2018 fifth international conference on social networks analysis, management and security (SNAMS)* IEEE, 336–341.

Sawan, S., Maalouf, R., Errachid, A., and Jaffrezic-Renault, N. (2020). Metal and metal oxide nanoparticles in the voltammetric detection of heavy metals: a review. *TrAC Trends Anal. Chem.* 131, 116014. doi:10.1016/j.trac.2020.116014

Trueman, B. F., Camara, E., and Gagnon, G. A. (2016). Evaluating the effects of full and partial lead service line replacement on lead levels in drinking water. *Environ. Sci. and Technol.* 50, 7389–7396. doi:10.1021/acs.est.6b01912

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. Neural Inf. Process. Syst.* doi:10.48550/arXiv.1706.03762

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. (2018). "Graph attention networks," in *International conference on learning representations (ICLR)*.

Vlachou, E., Margariti, A., Papaefstathiou, G. S., and Kokkinos, C. (2020). Voltammetric determination of pb(ii) by a ca-mof-modified carbon paste electrode integrated in a 3d-printed device. *Sensors* 20, 4442. doi:10.3390/s20164442

Vrahatis, A. G., Lazaros, K., and Kotsiantis, S. (2024). Graph attention networks: a comprehensive review of methods and applications. *Future Internet* 16, 318. doi:10.3390/fi16090318

WebHarvy (2022). What is web scraping.

Zanfei, A., Brentan, B. M., Menapace, A., Righetti, M., and Herrera, M. (2022). Graph convolutional recurrent neural networks for water demand forecasting. *Water Resour. Res.* 58, e2022WR032299. doi:10.1029/2022WR032299

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., et al. (2020). Graph neural networks: a review of methods and applications. *AI Open* 1, 57–81. doi:10.1016/j.aiopen.2021.01.001

Zillow (2023). Zillow: real estate, apartments, mortgages and home values.

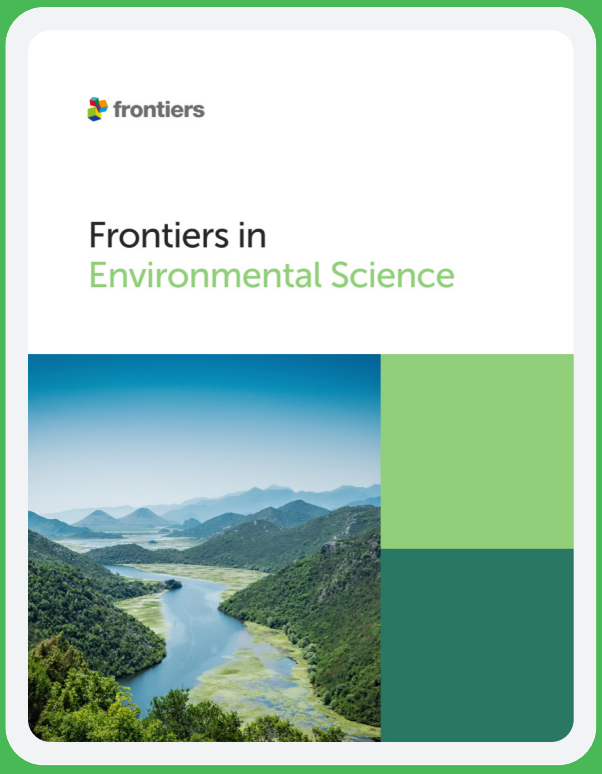
Frontiers in
Environmental Science

Explores the anthropogenic impact on our natural world

An innovative journal that advances knowledge of the natural world and its intersections with human society. It supports the formulation of policies that lead to a more inhabitable and sustainable world.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact