Information Fusion 114 (2025) 102667

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full length article

A Contemporary Survey on Multisource Information Fusion for Smart Sustainable Cities: Emerging Trends and Persistent Challenges

Houda Orchi [a,b,*], Abdoulaye Baniré Diallo [b], Halima Elbiaze [b], Essaid Sabir [c], Mohamed Sadik [a]

[a] NEST Research Group, LRI Lab, ENSEM, Hassan II University of Casablanca, Morocco
[b] Department of Computer Science, University of Quebec at Montreal, QC, H2L 2C4, Canada
[c] Department of Science and Technology, TÉLUQ, University of Quebec, Montreal, QC, H2S 3L4, Canada

ARTICLE INFO

Keywords:
Multisource information fusion
Hierarchical and multi-dimensional classification
Smart sustainable cities

ABSTRACT

The emergence of smart sustainable cities has unveiled a wealth of data sources, each contributing to a vast array of urban applications. At the heart of managing this plethora of data is multisource information fusion (MSIF), a sophisticated approach that not only improves the quality of data collected from myriad sources, including sensors, satellites, social media, and citizen-generated content, but also aids in generating actionable insights crucial for sustainable urban management. Unlike simple data fusion, MSIF excels in harmonizing disparate data sources, effectively navigating through their variability, potential conflicts, and the challenges posed by incomplete datasets. This capability is essential for ensuring the integrity and utility of information, which supports comprehensive insights into urban systems and effective planning. This survey combines hierarchical and multi-dimensional classification to examine how MSIF integrates and analyses diverse datasets, enhancing the operational efficiency and intelligence of urban environments. It highlights the most significant challenges and opportunities presented by MSIF in smart sustainable cities, particularly how it overcomes the limitations of existing approaches in scope and coverage.

By considering social, economic, and environmental factors, MSIF offers a multidisciplinary approach that is pivotal for advancing sustainable urban development. Recognized as an essential resource for academics and practitioners, this study promotes a new wave of MSIF innovations aimed at improving the cohesion, efficiency, and sustainability of smart cities.

1. Introduction

Global urbanization is intensifying, with over half of humanity now living in urban settings. This proportion is expected to exceed 70% by the year 2050 [1]. Such rapid expansion of urban areas introduces severe sustainability challenges, as these regions already account for roughly 75% of global consumption of essential resources, including fossil fuels, primary energy, water, and raw materials. Predictions indicate that the overall resource consumption could surge to an overwhelming 90 billion tonnes in the forthcoming decades [2]. These startling figures emphasize the imperative for effective sustainable urban planning and management to accommodate these increasing needs.

In response to these growing urban challenges, the paradigm of smart sustainable cities has emerged as a crucial driver in urban development, leveraging advanced technologies to meet current demands and criteria for ecological sustainability, economic stability, and social fairness, while also preparing for future challenges [3]. At the heart of this technological advancement is Multisource Information Fusion (MSFI) [4], which acts as a vital infrastructure for integrating diverse data sources. MSFI taps into various data streams such as satellite imagery, traffic sensors, air quality monitors, and water network analyses [5]. MSFI transforms raw data into actionable insights that not only enhance urban resource management and infrastructure but also foster more proactive and predictive approaches to city governance. By employing advanced AI and machine learning algorithms to analyze extensive data sets [6], MSFI supports essential urban operations like traffic management [7], pollution control [8], and the distribution of energy resources [9], which are critical for sustainable infrastructure development. Moreover, it enhances public safety by optimizing emergency response times and resource allocation [10], integrating data

* Corresponding author at: Department of Computer Science, University of Quebec at Montreal, QC, H2L 2C4, Canada.
E-mail addresses: orchi.houda@courrier.uqam.ca (H. Orchi), diallo.abdoulaye@uqam.ca (A.B. Diallo), elbiaze.halima@uqam.ca (H. Elbiaze), essaid.sabir@teluq.ca (E. Sabir), m.sadik@ensem.ac.ma (M. Sadik).

https://doi.org/10.1016/j.inffus.2024.102667
Received 2 June 2024; Received in revised form 27 August 2024; Accepted 2 September 2024
Available online 4 September 2024

from video surveillance and emergency communication systems [11]. Furthermore, MSFI is essential for precise environmental monitoring, crucial for managing air quality, noise levels, and water pollution, thus ensuring public health and urban well-being [12]. Smart sustainable cities are increasingly adopting MSFI to address critical issues such as waste management, pollution, energy consumption, and public safety [13]. This method promotes a more responsive and adaptable urban planning process, which is key to achieving sustainable development goals. For instance, waste management greatly benefits from MSFI through optimized collection routes and waste generation data analysis, enabling more efficient and cost-effective operations [14].

However, integrating information from diverse sources with MSFI presents challenges, particularly in terms of data diversity, privacy concerns, and scalability of solutions [15]. Managing data privacy [16] and protecting against cyberattacks are major concerns, as MSFI often involves handling large amounts of personal data. Also, integrating heterogeneous technologies and standardizing data formats [17] remain significant obstacles to fully unlocking MSFI's potential.

To move forward, cities must adopt regulatory frameworks and policies that promote effective collaboration between the public and private sectors [18] encouraging innovation and implementing MSFI solutions. These frameworks must also ensure that deployed technologies adhere to security, privacy, and transparency principles while being inclusive and equitable for all citizens [19]. By creating an adaptive regulatory environment, cities can facilitate the rapid deployment of new technologies while managing the risks associated with large-scale data fusion. Furthermore, to maximize the benefits of MSFI, cities need to invest in robust and secure data infrastructures. This includes developing advanced communication networks and establishing data centers capable of processing and securely storing large amounts of information [20]. These investments are crucial for enabling real-time data management and agile responses to changing urban dynamics. Education and training of municipal staff and stakeholders are also vital to ensure the effective adoption and use of MSFI. Training programs should be implemented to familiarize municipal employees with information fusion, data analysis, and cybersecurity practices. This education will help create a skilled workforce fully leveraging MSFI technologies to transform urban services [21]. In addition to these internal measures, cities must also seek to collaborate internationally to share best practices, innovations, and lessons learned in implementing MSFI [22]. Such collaboration can accelerate the development of smart sustainable cities worldwide, allowing municipalities to benefit from each other's experiences and successes while adapting solutions to specific local contexts.

Furthermore, cities must address challenges [23] related to digital equity to ensure that all citizens benefit from the advantages of MSFI. This includes improving access to high-speed internet, providing inclusive digital services, and implementing measures to combat the digital divide [24]. By ensuring equitable access to technologies, cities can prevent the creation of new inequalities and strengthen social cohesion.

In conclusion, Multisource Information Fusion transcends mere technological advancement; it signifies a paradigm shift in urban governance, necessitating meticulous planning, cross-sector collaboration, and unwavering dedication to inclusivity and sustainability. Embracing this data-driven approach not only tackles present urban hurdles but also paves the way for constructing resilient and sustainable cities, primed to thrive in the years ahead.

The existing literature lacks comprehensive reviews that merge the latest advancements in MSIF techniques with their applications in sustainable smart cities. Motivated by this gap, our survey investigates pioneering and established MSIF methodologies that enhance operational efficiency and decision-making. We focus on recent innovations like advanced machine learning algorithms for predictive analytics, which have significantly transformed urban data management. By documenting and evaluating these state-of-the-art technologies, our main objective is to thoroughly analyze the architectures, challenges, and impacts of MSIF in sustainable smart cities.

Our main contributions in this paper are three-fold:

- We propose a hierarchical and multi-dimensional approach to examine the role of MSIF in sustainable smart cities, highlighting the challenges and opportunities encountered.
- We provide insights into the most effective MSIF strategies and their implications for sustainable urban development.
- We discuss the challenges and ideal scenarios for deploying MSIF techniques in sustainable smart cities, serving as a resource to inspire innovations that enhance integration, efficiency, and sustainability.

As shown in Fig. 1, the rest of this paper is organized on the following lines: Section 2 provides background information on MSIF and its evolution within sustainable smart cities, highlighting the methodologies utilized in our comprehensive study. Section 3 delves into various MSIF techniques and their applications, showcasing how these techniques enhance urban management. It covers models such as the JDL Model and Bowman Data Fusion Framework, along with theoretical aspects like traditional fusion theory approaches.

Section 4 outlines the theoretical foundations that underpin MSIF, focusing on our Advanced Hierarchical and Multi-Dimensional Classification Framework within urban data science. Section 5 identifies the urban challenges tackled by MSIF, proposing innovative solutions and addressing potential hurdles. This section encompasses data challenges, application challenges, and implementation challenges in MSIF for smart sustainable cities. Section 6 presents practical applications of MSIF through detailed case studies. These case studies illustrate the implementation and impact of MSIF in real-world scenarios, including environmental monitoring in Sichuan Province, smart grid integration in urban environments, and urban carrying capacity analysis in Shanghai.

Section 7 examines the prospects and research directions for MSIF, highlighting ongoing innovations and potential advancements for smart sustainable cities. This includes the integration of advanced AI, autonomous intelligent fusion, integrated cognitive systems, interdisciplinary collaboration, and exploration of novel research domains. Lastly, Section 8 concludes by reaffirming the transformative impact of MSIF on urban governance, emphasizing its indispensable role in promoting the development of sustainable and resilient smart cities.

2. Historical and methodological exploration

2.1. Background

The terminology 'data fusion' and 'sensor fusion' first emerged in the literature [25] during the late 1970s. Despite its early mentions, a cohesive understanding of information fusion technologies did not materialize until the 1990s when the term 'information fusion' began to be widely accepted. This period marked the establishment of Multisource Information Fusion (MSIF) as a recognized interdisciplinary field. Following this, MSIF technologies experienced rapid advancements, leading to significant applications both in military contexts, such as in automated systems [26], strategic early warning and defense systems [27], target tracking [28], and damage assessments [29] and in civilian domains, including remote sensing [30], healthcare diagnostics [31], e-commerce [32], telecommunications [33], navigational aids [34], and fault diagnosis [35].

Despite decades of advancement and application, a standardized definition of MSIF remains elusive in academic and industry sectors. From a military standpoint, the Joint Directors of Laboratories (JDL) of the United States Military first offered a formal definition, known as Definition 1: MSIF is an extensive, layered methodology that entails the detection, association, integration, and analysis of data from diverse sources. This approach enhances the precision of state and identity estimation while ensuring a timely and thorough assessment of significance. This foundational definition was expanded upon in Definition 2, often referred to as multi-sensor fusion, which characterizes MSIF

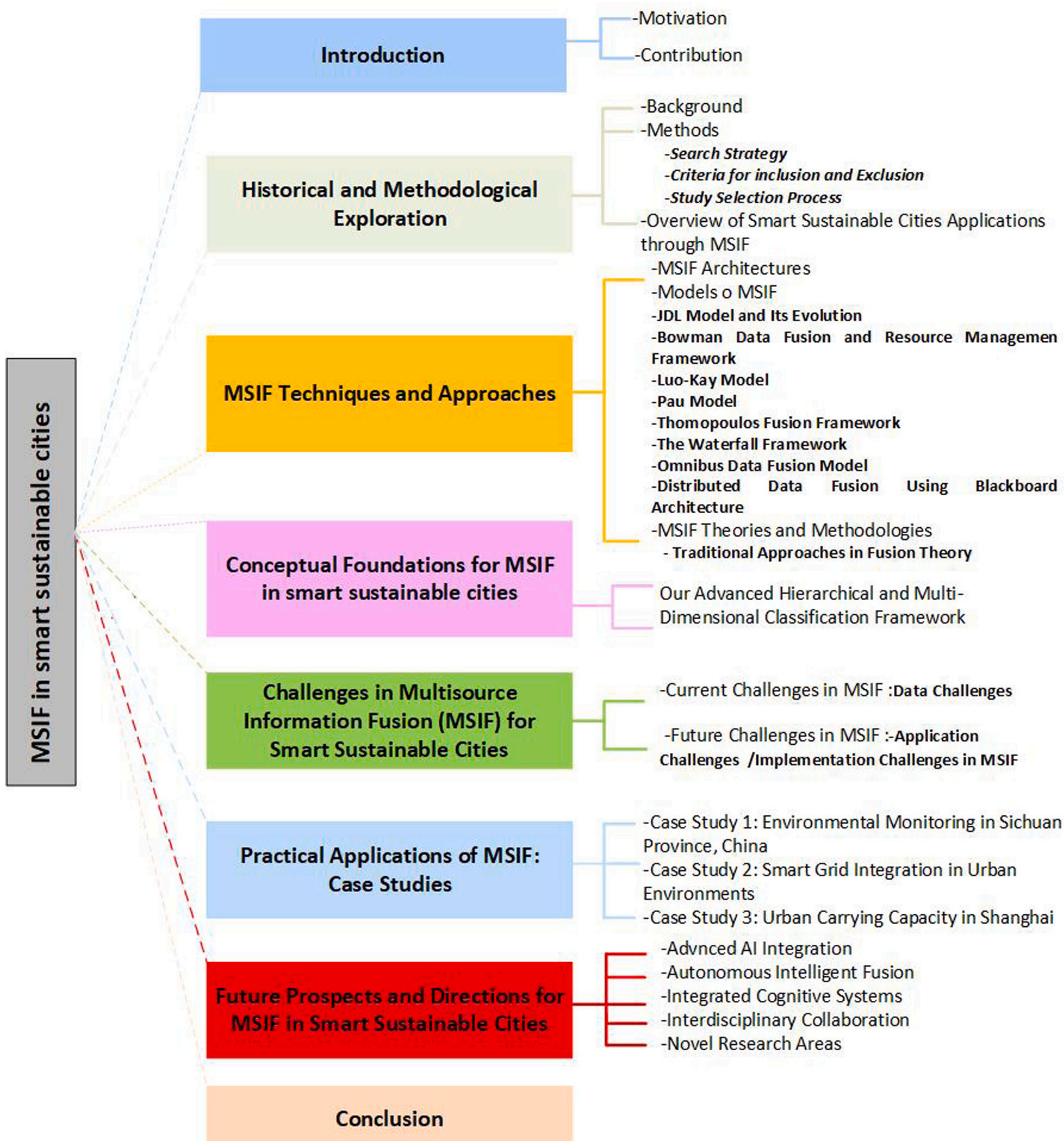

Fig. 1. Roadmap and core content of this article.

as a multi-layered and complex procedure encompassing the detection, association, integration, and evaluation of data from various sources. This process aims to enhance the accuracy of information, state, and identity estimation, as well as to provide a timely and comprehensive assessment of the overall target situation and threat level.

Definition 3 positions MSIF as an effective method for automatically or semi-automatically transforming information from different sources and time points into representations that provide effective support for human or automated decision-making.

Further, Definition 4 elaborates that MSIF is a process of detecting, representing, mining, associating, synthesizing, reasoning, and predicting multi-source, multi-level, multi-dimensional, and granular data, information, and knowledge. This process integrates modern information technologies, such as Artificial Intelligence (AI), knowledge reasoning, data analysis, and processing, to obtain richer, more accurate, and consistent descriptions or decision results concerning the observed object.

2.2. Methods

This study adhered to rigorous systematic review and meta-analysis protocols to provide a meticulous and in-depth analysis. A comprehensive search was conducted across various key bibliographic databases such as ScienceDirect, Scopus, IEEE Xplore, arXiv, and Hindawi to ensure broad coverage of relevant research, providing valuable insights into multisource information fusion in smart sustainable cities.

2.2.1. Search strategy

From 2010 through April 2024, research from databases like ScienceDirect, arXiv, IEEE Xplore, Scopus, and Hindawi was scrutinized. The focus was on studies related to MSIF in smart sustainable cities, using a Boolean search with terms like "multisource information fusion" and "smart sustainable cities" to collect a wide range of pertinent studies. Significant foundational works, even those published before 2010 that are crucial to the field – such as foundational theories in information fusion from established articles and well-known frameworks like JDL and Dasarathy – were also considered in the review.

2.2.2. Criteria for inclusion and exclusion

The inclusion criteria specified:
-Articles must be written in English and published in respected academic journals or at conference proceedings.
-Research must be specifically related to multisource information fusion within the context of smart sustainable cities.
-The studies should substantially deepen the comprehension of multisource information fusion in these urban environments by leveraging different information fusion techniques, architectures and hybrid frameworks.
The exclusion criteria ruled out:
- Non-English articles.
-Studies that only slightly touch upon data fusion in smart cities.
- Reviews and empirical studies that are narrowly focused on particular hypotheses without contributing significant insights or innovative proposals.

2.2.3. Study selection process

Following PRISMA guidelines as shown in Fig. 2, the study employed a methodical selection approach. Initially, duplicate entries were eliminated using Mendeley software. This was followed by scrutinizing titles and abstracts to determine their relevance. Initially, 330 articles were identified, reduced to 312 after duplicates were removed. Detailed scrutiny of titles and abstracts reduced this number to 253 for full-text review, ultimately refining down to 231 articles that provided a thorough understanding of the research topic.

2.3. Overview of smart sustainable cities applications through multisource information fusion

This section explores the applications of multisource information fusion (MSIF) in smart sustainable cities, highlighting its role across various urban domains. MSIF integrates diverse data sources to provide a holistic evaluation of urban systems, thereby enhancing their efficiency and sustainability. By merging data from multiple sources, MSIF improves living standards, facilitates urban management, and promotes environmental sustainability.

Smart Living Applications

Smart living focuses on enhancing urban quality of life through the integration of MSIF technologies in healthcare, home automation, and community services. This includes monitoring health conditions using sensors and data fusion to offer comprehensive healthcare solutions, automating homes with intelligent systems that learn from user behaviors and environmental data, and creating connected community networks that enhance public engagement and service delivery.

Urban Area Management

MSIF applications in urban management involve the use of data from various sources for efficient city governance and infrastructure management. This includes systems for intelligent building management that adapt to usage patterns, fire detection systems that integrate data from multiple sensors to improve safety, and comprehensive urban planning tools that use MSIF to analyze extensive datasets for better urban development strategies.

Environmental Sustainability

In the environmental sector, MSIF helps manage resources and monitor urban environments effectively. Applications include pollution monitoring systems that consolidate data from multiple air quality sensors, waste management systems that optimize routes and schedules based on real-time data, and water quality management that uses sensors across water bodies to ensure safety and sustainability.

Industrial Applications

Industrial applications of MSIF enhance operational efficiency and sustainability in manufacturing and agriculture. This includes predictive maintenance systems that use data from various machine sensors to foresee and prevent breakdowns, and precision agriculture systems that integrate soil, weather, and crop data to optimize farming practices and resource use.

Economic Development

Economic applications of MSIF in smart cities enhance the efficiency of commercial activities through better data integration. This encompasses supply chain management systems that use real-time data fusion to streamline operations, dynamic pricing models based on multisource market data, and enhanced customer relationship management systems that integrate consumer data across multiple platforms.

Human Mobility and Transport

MSIF improves transportation systems by integrating data from various transportation modes and traffic sensors to manage traffic flow, optimize public transport, and provide real-time navigation aids, thereby enhancing mobility and reducing congestion in urban areas.

Infrastructure Optimization

This sector uses MSIF to manage and optimize critical urban infrastructure like power grids, water systems, and public facilities by integrating data from various infrastructure sensors to predict maintenance needs, enhance resource allocation, and ensure reliable service delivery.

These applications showcase how multisource information fusion can significantly enhance the functionality, sustainability, and livability of smart sustainable cities across multiple domains. To better illustrate these concepts, we present a figure depicting the layered infrastructure of smart sustainable cities. As shown in Fig. 3, the infrastructure of smart sustainable cities consists of several key layers:

-Sensing Layer: This layer consists of various sensors deployed throughout the city to collect data on different aspects such as air quality, traffic flow, and energy usage.
-Transmission Layer: Responsible for transmitting the collected data to central processing units using wired or wireless communication networks.
-Blockchain Layer: Ensures secure and transparent data transactions across the network, enhancing data integrity and trust.
-Data Management Layer: Involves storing, processing, and analyzing the data to extract valuable insights.
- Smart Application Layer: Utilizes the processed data to develop intelligent applications for healthcare, transportation, environment monitoring, and more.

This layered approach ensures that smart sustainable cities can effectively utilize MSIF to improve urban living standards and sustainability.

3. MSIF techniques and approaches

Multisource Information Fusion (MSIF) integrates data from various sources to create comprehensive, accurate, and actionable information. MSIF techniques are invaluable in smart sustainable cities, where diverse data streams require intelligent and efficient integration to support decision-making. This section explores the primary architectures, models, theories and approaches used in MSIF, highlighting

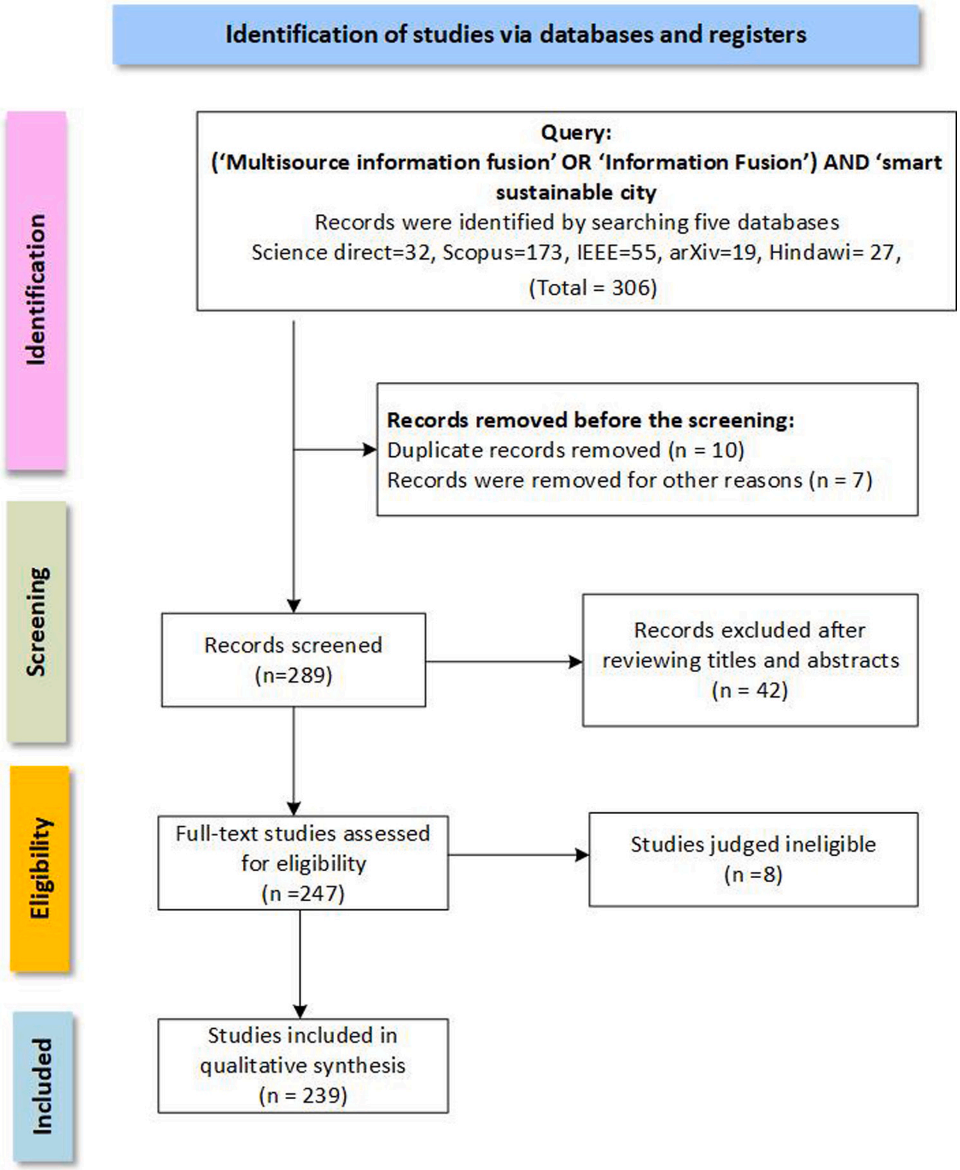

Fig. 2. A diagram illustrating the identification, screening, and inclusion of relevant studies.

their unique characteristics and how they can be leveraged in urban environments. Currently, there are many widely recognized data fusion architectures and classification standards. These frameworks offer structured approaches for integrating data from multiple sources. Fig. 4 illustrates different architectures and classification standards for data fusion technology. To delve deeper into the specific types of fusion architectures, Fig. 5 categorizes them into three main groups: data-level, feature-level, and decision-level fusion. This classification underscores the range and depth of fusion approaches available in MSIF, showcasing how each category plays a crucial role in the context of smart sustainable cities.

3.1. MSIF architectures

MSIF draws inspiration from various fields, including signal processing, information theory, statistical estimation and inference, and AI. Traditional fusion methods are primarily based on statistical inference and estimation theory, providing a solid theoretical foundation for MSIF. Recent advancements in AI and information theory have led to significant progress in MSIF techniques. To better understand these advancements, Table 1 offers insights into key methodologies, challenges, and technological developments within MSIF. The methods are categorized into three main types, each defined by the level at which information sources are processed:

- *Data-Level Fusion:* [43] The lowest level fusion combines raw data from different sources to create a unified dataset that can be processed holistically. The attribute decision at this stage involves determining how to efficiently combine these data streams, considering factors like data alignment and normalization to reduce preprocessing needs. This approach is most effective when dealing with homogeneous data or when data pre-processing is minimal. In sustainable smart cities, data-level fusion can be beneficial in real-time traffic management systems that aggregate data from traffic cameras, GPS devices, and traffic signal controllers to offer a live overview of the traffic flow.
- *Feature-Level Fusion*: [44] Feature-level fusion extracts relevant features from individual data sources before merging them into a single feature vector. The attribute decision here focuses on the selection and integration of the most informative features, ensuring that the combined feature vector maximally represents the underlying data for subsequent analysis. This technique provides a more abstract representation of the data, allowing advanced machine learning algorithms to

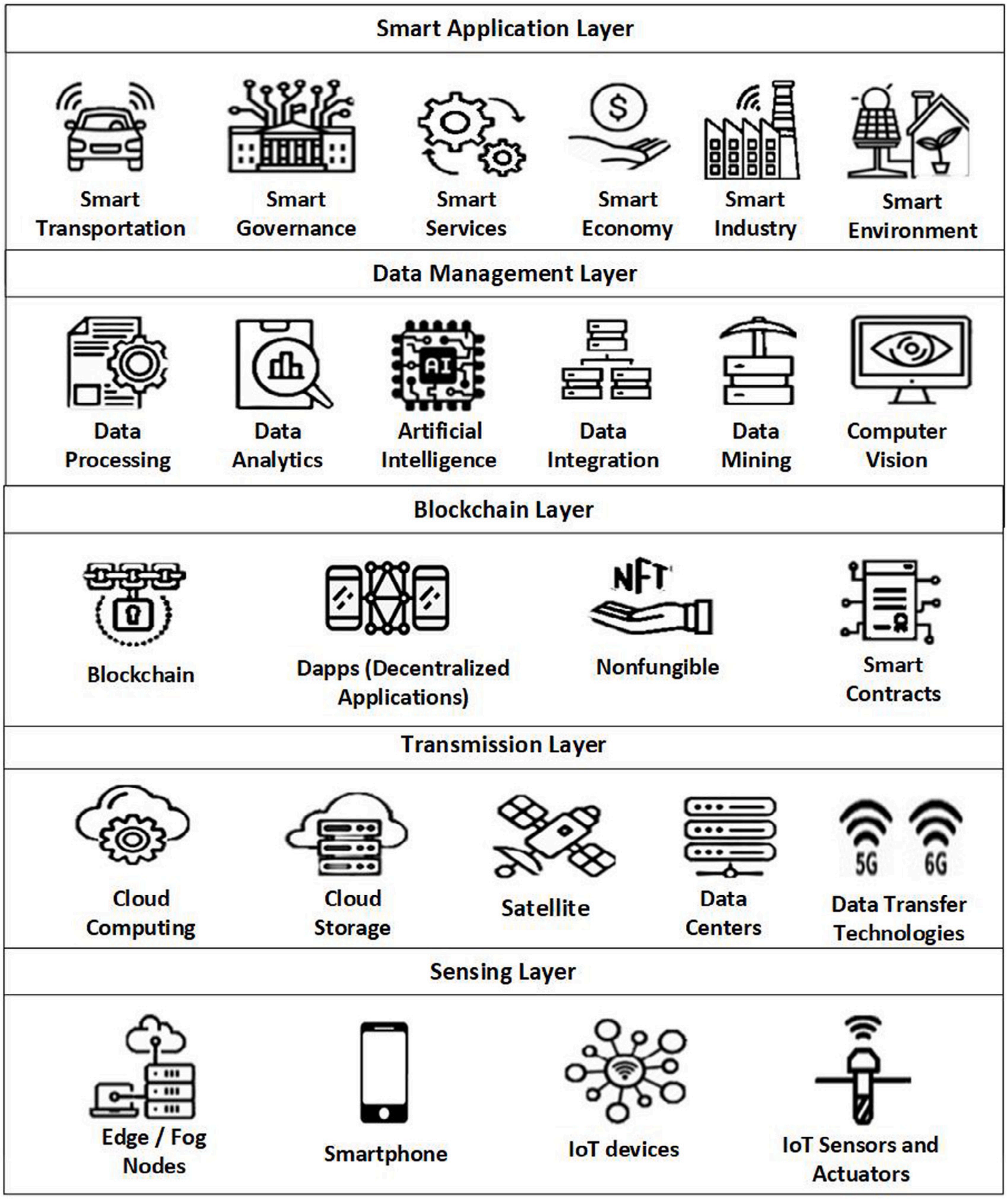

Fig. 3. Smart sustainable city layered infrastructure.

find correlations that would otherwise remain undetected. For example, feature-level fusion can be used to analyze environmental data from air quality sensors and weather stations to develop predictive models for air pollution.

- *Decision-Level Fusion*: [45] Decision-level fusion consolidates the results of independent classifiers or decision-makers to reach a final decision. The attribute decision in this context involves choosing the optimal way to combine these independent results to leverage their strengths and mitigate any weaknesses, thus enhancing the reliability of the final decision. This approach is useful when individual sources provide distinct yet complementary insights into a problem. In the context of smart cities, decision-level fusion can combine traffic incident reports, weather forecasts, and emergency response data to coordinate rescue efforts during natural disasters. In addition to these levels, other fusion methods further enhance the accuracy and effectiveness of the MSIF process:

- *Hybrid Fusion*: [46] Hybrid fusion combines two or more of the techniques above to improve the overall accuracy and robustness of the fusion process. It leverages the strengths of different fusion levels to provide comprehensive and accurate insights. For instance, a hybrid approach can first apply feature-level fusion to aggregate data before employing decision-level fusion to refine predictions made by multiple models. In the healthcare sector, hybrid fusion could assist in the predictive modeling of disease outbreaks.

- *Context-Aware Fusion*: [47] Context-aware fusion incorporates contextual information into the fusion process, enhancing decision-making by understanding the specific circumstances under which the data is collected. For example, integrating data from social media, news outlets, and IoT devices with contextual information about urban demographics allows better modeling of emergency responses in densely populated urban areas.

- *Machine Learning-Based Fusion*: [48] Machine learning-based fusion relies on algorithms that can learn the optimal ways to combine data

Table 1
Comparison of MSIF surveys in terms of methodologies, challenges, and technological advancements.

References	Objectives of MSIF	Fusion techniques	Data types	Challenges addressed	Case studies	Emerging trends	Technological innovations	Future directions	Gaps in the study
[36]	–	Data association T1, state estimation T2, decision fusion T3, and classification T4.	- Sensor data, IoT data, big data.	- Data quality, data representation, data privacy and security.	–	–	- Communication technologies (5G, WSN, LPWAN, NB-IoT) and data mining techniques.	- Improving data quality, enhancing data representation, and ensuring data privacy and security.	- The need for enhanced sensing coverage and longevity to ensure data quality, addressing the challenges of data representation for effective integration.
[37]	–	- Fuzzy reasoning-based fusion,PredRNN and GC-PredRNN-based fusion, hybrid fusion (fusion based on convolutional neural networks, decision-based (voting) fusion, Bayesian fusion, Dempster–Shafer fusion), Kalman filtering and learning-based fusion.	- Environmental, transportation, social, infrastructural data, sensor data, geographic information	- Data heterogeneity, quality control, privacy, scalability, interpretability of ML models.	- Traffic management, environmental monitoring, urban livability, emergency response.	- IoT, big data, data mining, information fusion, sustainable urban development.	–	- Enhancing urban management practices, optimizing resource allocation, improving quality of life.	-A scarcity of literature on critical areas such as Cyber–Physical Systems for disaster prevention, traffic management systems, and the role of mobile intelligent actors in social computing. -The integration and interoperability of diverse data sources and systems.
[38]	–	- Sensor fusion, feature-based data fusion, decision fusion.	- Physical data sources, cyber data sources, participatory sources, hybrid data sources.	–	-Crowd monitoring, automation applications.	–	- synchronization of temporal algorithms, data integrity, data security, confidentiality mechanisms.	–	-The need for holistic data fusion architectures, advanced fusion technologies, and in-depth research on data fusion techniques at every stage of the data value chain.
[39]	–	- Probabilistic-based methods (e.g., Kalman filter and its variations), Evidence reasoning-based methods, (e.g., Dempster–Shafer theory).	- Heterogeneous data types include location, images, speed, flow, weather, and sensor data.	- Data missing, redundancy, delay, anomalies, datasets normalization, and sensor data translation into formal languages	- Various real-life environments and simulations across different ITS applications such as fatigue detection, flow prediction, bandwidth allocation, tracking, and 3D object detection.	- Multi-sensor data sources, hierarchical fusion structures, adaptive and federated Kalman filter fusion models, and integration of time-varying information.	–	- Strengthening safety and reliability in ITS, improving robustness and accuracy, and reducing potential road accidents.	-Research gaps include the need for more efficient data processing, fault-tolerance, and the handling of incorrect data from sensors.
[40]	–	- Methods for data fusion include aligning input data both temporally and spatially, associating data, and employing data mining techniques.	- Various data sources including real-time point and wide-area traffic flow sensors, commercial vehicle transmissions, and roadway-based weather sensors.	- Feasibility, effectiveness, and usefulness assessment of DF approaches accuracy, dynamic and real-time aspects of traffic, and data quality.	–	- Integration of wireless technologies for easier reporting and access to customized information, and enriching available information on traffic situation.	–	- Increased collection of usable data from different sources, evaluation of benefits of DF through successful practical applications.	-Accuracy necessary for effective application, real-time dimension, and assessment of benefits through practical applications.
[41]	–	-The integration of the data collected from these IoT and sensor-based sources with other data systems, such as geographic information systems (GIS).	- Product lifecycle data. -Service information.	- Limited case studies on SC sustainability. -The need for on-time waste collection, separation, and integration of different waste management systems. - Real-world validation of the framework.	- Smart cities in Spain. -Energy project in Bolzano, Italy.	- Circular economy. - The sharing economy. - Industry 4.0.	- Cyber–physical systems (CPS), blockchain technology, and Internet of Things (IoT).	- Consideration of regulation, policy, product design strategies, and technology. - Tuning the framework for different waste types. - Validation with real-world case studies.	- Different waste types require specific management systems. - Infrastructure for data collection needs further study. - Political, legal, and commercial barriers.
[42]	- To improve the efficiency of performance management of power enterprises through the MSIF model.	-Data-level fusion. -Model-level fusion. -Decision-level fusion.	- Level fusion. -Electrical quantity. -Process quantity. -State quantity. -Wind speed. -Reactive power. -Cabin and outdoor temperature.	- Redundancy and noise in initial data, ensuring data quality and effectiveness, preprocessing and classification of data.	–	Big data analytics for smart manufacturing, integration of big data into corporate decision-making.	- Algorithms for big data processing (e.g., MP-Hermite, BPNN), visualization techniques for complex data.	- Refinement of performance evaluation indexes, more comprehensive reflection of enterprise performance evaluation.	- Not refined enough performance evaluation indexes, partially reflecting enterprise performance information.
Our contemporary survey	- To revolutionize MSIF strategies for enhanced urban planning and sustainability through the introduction of our Advanced Hierarchical and Multi-Dimensional Classification Framework.	- Hierarchical and multi-dimensional classification, integration of real-time and historical data fusion, advanced AI-driven analytics.	- Extensive urban datasets including traffic sensors, satellite imagery, social media content, citizen-generated data, environmental sensors, and administrative records.	- Complex data integration, scalability across diverse urban environments, ensuring data security, enhancing privacy protections, and improving interoperability among heterogeneous data systems.	- Detailed case studies on urban traffic management, resource distribution, environmental monitoring, and public safety improvements.	- Edge computing for faster data processing, predictive analytics for proactive urban management, AI for autonomous decision-making, IoT integration with urban infrastructures.	- Deployment of neural networks for pattern detection and analysis, use of blockchain for secure data transactions and smart contracts for maintaining data integrity, and advanced algorithms for efficient real-time data processing.	- Expanding the framework to accommodate emerging technologies like quantum computing for data processing and AI advancements for smarter analytics. Developing interoperable standards to facilitate broader integration of MSIF across various platforms and international boundaries.	- Our survey fills significant gaps by addressing the need for comprehensive data fusion architecture and practical, scalable integration strategies.

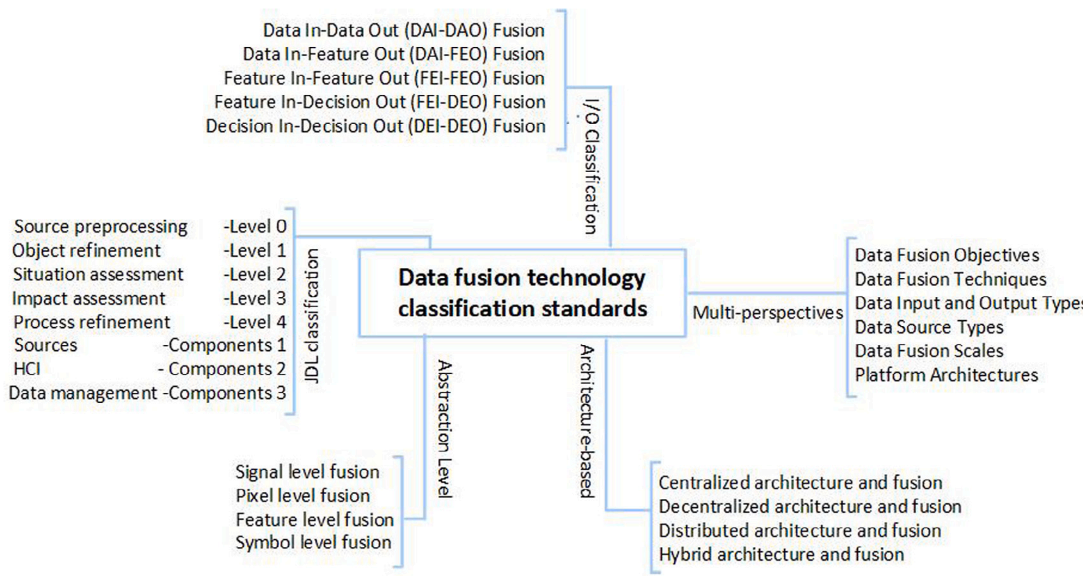

Fig. 4. Various data fusion architectures and technology categorization standards.

from different sources. Deep learning models, in particular, excel at discovering intricate patterns in multisource datasets. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [49] have been successfully applied to analyze imagery, sensor, and textual data simultaneously to support comprehensive decision-making in urban planning.

- *Bayesian Fusion*: [50] Bayesian fusion uses probabilistic models to merge data based on conditional probabilities. It is particularly effective in situations where data sources have varying levels of reliability. Bayesian networks or belief networks model the relationships among various data sources, providing a more realistic representation of uncertainty. This technique is widely used in security systems for anomaly detection and surveillance in public spaces.

3.2. Models of MSIF

3.2.1. JDL model and its evolution

The Joint Directors of Laboratories (JDL) model is structured around data fusion processes, including correlation, filtering, and association. This model organizes data fusion into five sequential stages: source preprocessing (level-0), object refinement (level-1), situation refinement (level-2), threat refinement (level-3), and process refinement (level-4), as depicted in Fig. 6. The model is designed so that data progresses from data sources to a human–computer interface, making the utilization of contextual information essential for advancing from low-level to high-level fusion tasks. In 1998, the Data Fusion Information Group (DFIG) model was introduced as an important enhancement to the JDL model to better incorporate contextual information. The DFIG model added capabilities for data fusion and resource management, emphasizing user interaction as observers to help contextualize during the data fusion process [51]. A further revision in 2004 led to the reassessment of the DFIG-enhanced JDL model [52], pinpointing several shortcomings such as the need for users to apply varied skills – perception, task rules, and cognition – to manage different tasks, with many systems tailored only to specific requirements. Post-2004, the focus shifted towards enhancing information management, advanced visualization, data mining, and the integration of teamwork, prioritization, and coordination in fusion system design. These enhancements aimed to tackle the challenges of fusion and resource management within the DFIG framework.

Subsequent expansions in 1998 and 2004 addressed several of the JDL model's initial limitations, which overly emphasized data (input/output) at the expense of processing and faced restrictive constraints. Despite these enhancements, practical applications of the JDL model still encounter challenges. In response, the Dasarathy model [53] was developed to view fusion systems from a software engineering perspective, treating them as data flows characterized by inputs/outputs and functions. This model is primarily utilized for optimal decision-making in multi-sensor target recognition and tracking, seeking to fuse data from a network of sensors in parallel within a recursive system structure to enhance decision reliability [54]. Addressing decision reliability further, Goodman and al. [55] proposed a fusion model based on random sets, integrating decision uncertainty directly into the fusion process and offering a comprehensive framework for representing uncertainty.

3.2.2. Bowman data fusion and resource management framework

The Bowman Data Fusion&Resource Management (Df&Rm) model, introduced by Bowman [56] in 1980, presents a versatile data fusion architecture aimed at resolving issues related to multi-sensor, multi-target identification, and tracking. Although the JDL model has been effective in various data fusion applications, Bowman observed that it had limited influence on the architectural development of practical systems. As a result, Bowman proposed the concept of a data fusion hierarchy tree, which breaks down fusion problems into nodes. Each node conceptually includes functions such as data linking, estimation, and association. In the Bowman Df&Rm model, data processing is based on correlation assumptions and prior information to evaluate the impact of uncertain and unknown targets on data and decision outcomes. The model incorporates a structure that includes hypothesis generation and evaluation feedback (hypothesis validation), as illustrated in Fig. 7. Additionally, the Bowman Df&Rm model highlights the duality between estimation and control, where data fusion and resource management systems are achieved through the combination and management of interactions between network nodes. This interaction interprets various data types, sources, models, and conclusions distinct within the Bowman Df&Rm architecture. Within this model, the process

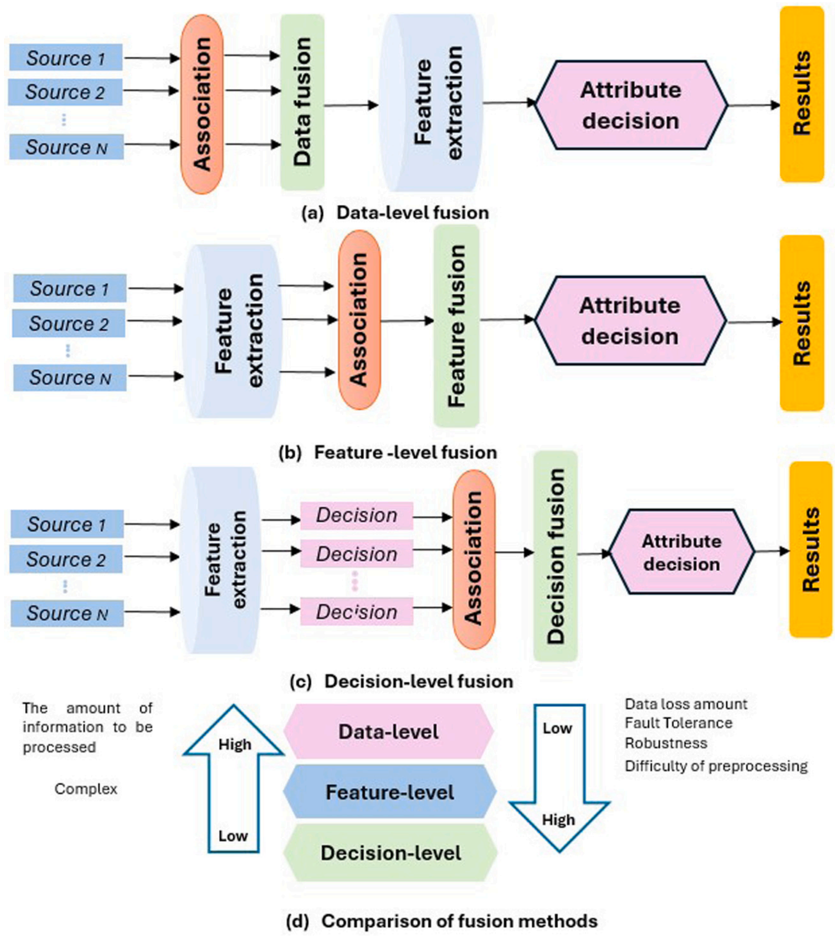

Fig. 5. Three categories of fusion architectures and their comparison.

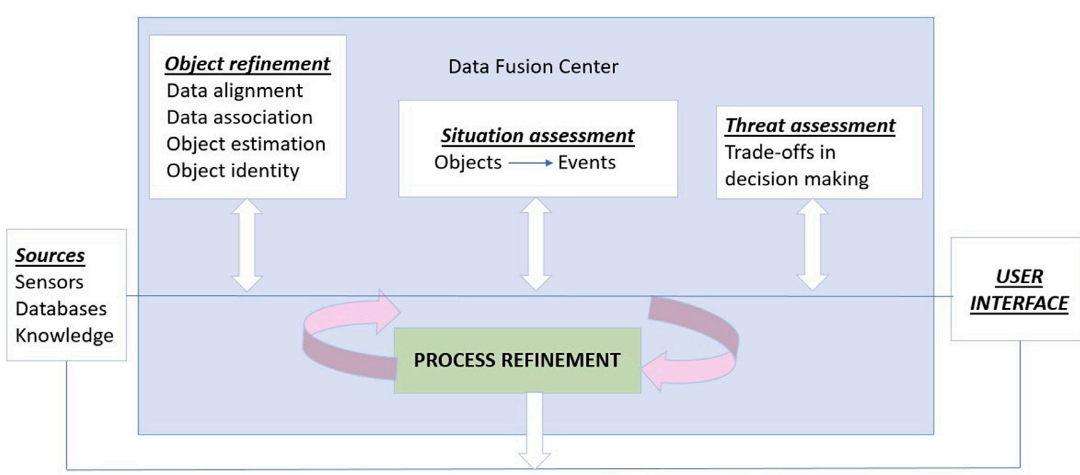

Fig. 6. The JDL framework for data fusion.

refinement function considered a component of resource management functions, corresponds to the fourth level in the JDL model as revised in 2004.

3.2.3. Luo-Kay model

Luo and Kay [57] proposed a versatile data fusion framework centered on multi-sensor integration. In this framework, data from multiple sources is systematically combined within embedded fusion centers arranged hierarchically. They distinguished between multi-sensor integration, which leverages information from multiple sensors to support specific tasks, and multi-sensor fusion, which involves the actual combination of data at various integration stages. Fig. 8 delineates Luo and Kay's structured approach to multi-sensor integration and fusion. In this framework, sensor data is systematically relayed to hierarchical fusion centers, where it undergoes a layered fusion process. This methodical approach is pivotal in synthesizing data from

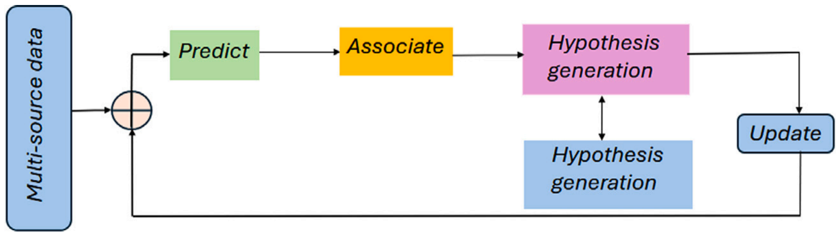

Fig. 7. Bowman data fusion and resource management framework.

Table 2
Properties of data fusion layers.

Characteristics	Signal level	Pixel level	Feature level	Symbol level
Information representation level	Low	Low	Medium	High
Type of sensor data	Multidimensional signals	Multiple images	Extracted features from signals/images	Decision logic from signals/images
Sensor data model	Noisy random variable	Random process across pixels	Non-invariant feature forms	Symbols with degrees

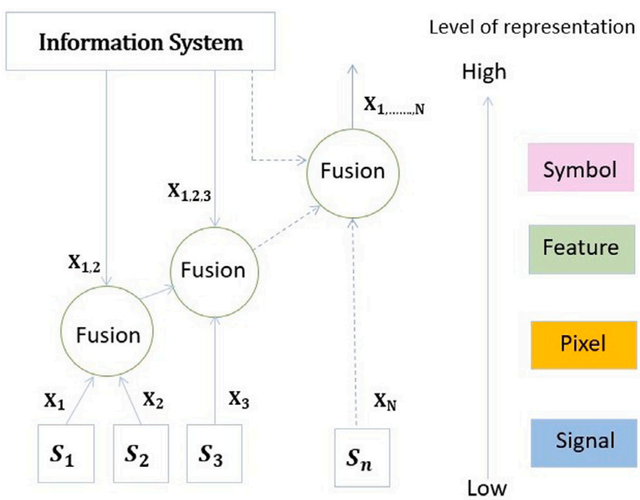

Fig. 8. Luo and Kay's structural model.

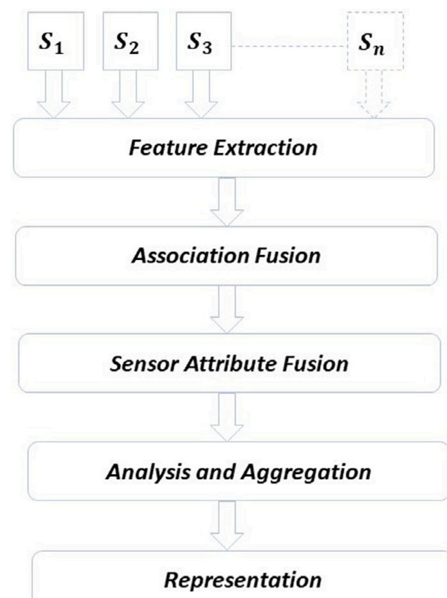

Fig. 9. Pau's approach for fusing sensor data.

various sensors to generate a detailed understanding of observed phenomena. The Information System, integral to this process, incorporates comprehensive databases and sophisticated algorithms to manage data flow and enhance quality at each stage of integration. Within the Information System, advanced algorithms are employed not only to ensure the precision and coherence of the data but also to maintain stringent security protocols that safeguard data integrity and adhere to regulatory standards. As data progresses through the fusion centers, it is transformed from raw signals into progressively more sophisticated and abstract forms, culminating in symbolic representations that provide actionable insights.

The progression of data through the Information System is crucial for effectively handling and interpreting diverse data types, ensuring comprehensive integration and meaningful analysis. Table 2 in the manuscript provides a systematic comparison of these fusion levels, detailing how each contributes to the overall functionality and effectiveness of Luo and Kay's model, emphasizing the seamless transition from raw data collection to advanced information synthesis.

3.2.4. Pau model

The Pau model, characterized by a typical hierarchical architecture, employs a data fusion approach grounded in behavioral knowledge [58]. Initially, feature vectors are extracted from raw data and aligned with defined attributes. These vectors are then combined and analyzed at the sensor characteristic fusion and data analysis levels, culminating in a decision stage where behavioral rules are applied based on explicit combination outputs. The Pau model is depicted in Fig. 9 and consists of three display levels, starting with each sensor having a vector space that includes coordinates and measurement parameters, followed by feature extraction and labeling, and culminating in the association of feature vectors with events and the definition of environmental models and fusion strategies.

3.2.5. Thomopoulos fusion framework

Thomopoulos [59] proposed a data fusion architecture composed of three distinct modules, each responsible for integrating data at different levels:

– Signal Level Fusion: This level involves data correlation achieved through learning processes, primarily because no mathematical model is available to describe the measured phenomenon.

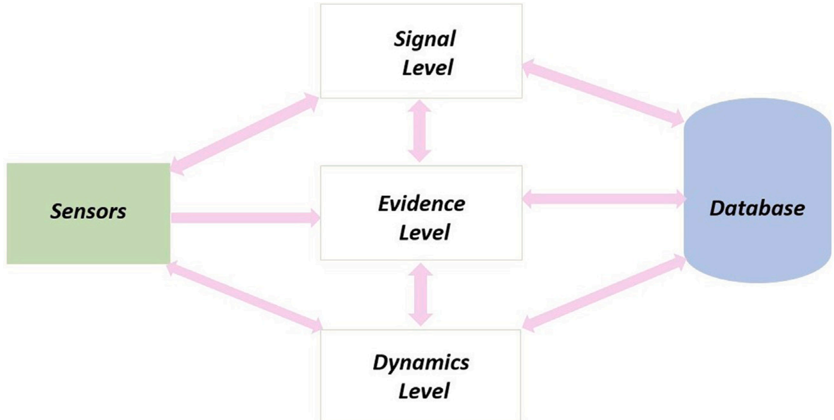

Fig. 10. Thomopoulos's framework for fusing data.

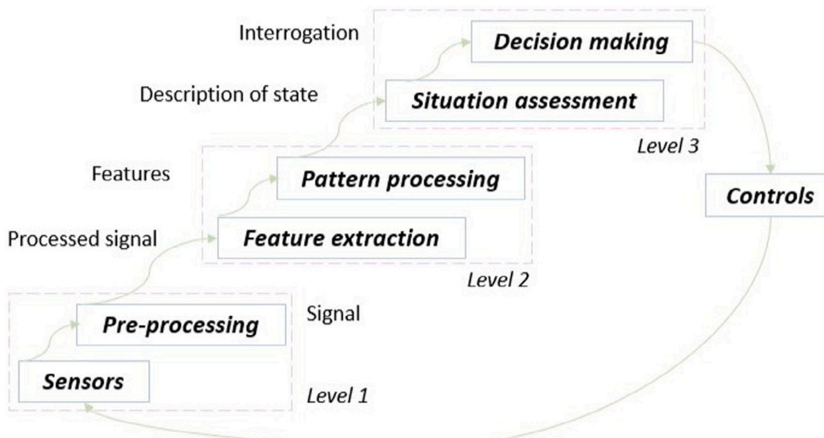

Fig. 11. The waterfall framework for sequential progression.

– Evidence Level Fusion: At this stage, data is combined at various inference levels using a statistical model. This level is tailored to meet the user's requirements, such as decision-making or hypothesis testing.
– Dynamics Level Fusion: Here, data fusion is performed using an existing mathematical model to guide the process.

These levels of fusion can be implemented either sequentially or interchangeably, depending on the application. For objectives such as continuous health monitoring of machinery, data fusion can initially occur at the signal level. However, for making comprehensive decisions, a higher level of data integration, such as evidence fusion, is required. This approach is essential when decisions must be derived from a broad analysis of the signals. Fig. 10 illustrates the overall architecture of this process.

Thomopoulos highlights that for a data fusion system to function effectively, it must adhere to three essential principles:

– Information Monotonicity: As more information is added, the system's performance should improve.
– Cost Efficiency: The costs should increase proportionally with the amount of data processed and integrated.
– Resilience to Pre-existing Uncertainties: The system needs to effectively manage initial uncertainties present before the fusion of data.

In addition, operational considerations such as data transmission delays, communication errors, and the accurate alignment of data in both space and time are crucial for the data fusion system's effective operation. These factors significantly influence the system's performance and dependability.

3.2.6. The waterfall framework

Harris [60] introduced a hierarchical architecture frequently adopted by the data fusion community, known as the waterfall model. This model is depicted in Fig. 11, illustrating the flow of data from the initial data level to the final decision-making level. The sensor system receives continuous updates via feedback from the decision-making module. This feedback informs the multi-sensor system about necessary recalibration, reconfiguration, and data collection.
The waterfall model consists of three representation levels:
Level 1: Raw data is transformed to provide essential environmental information. To perform this transformation, models of the sensors and, whenever possible, models of the phenomena being measured are utilized. These models may be derived from experimental analysis or physical laws.
Level 2: This level involves feature extraction and the fusion of these features. The goal is to derive symbolic inferences from the data, minimizing the amount of data while maximizing the information content. The output is a list of estimates, each with associated probabilities and confidence levels.
Level 3: This level associates objects with events, compiling possible actions based on the gathered information, available libraries and databases, and human input.

Below is a visual representation of the waterfall model: This hierarchical approach enables a structured flow of data processing, ensuring

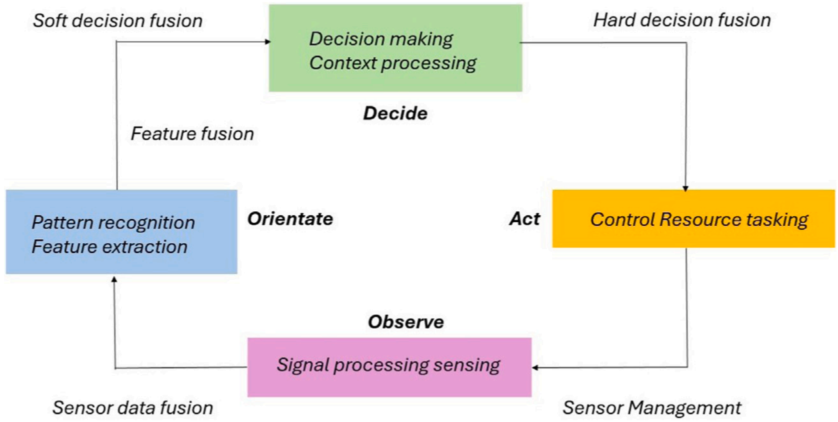

Fig. 12. The omnibus framework for fusing data.

that each stage refines and improves the information before reaching the decision-making phase.

3.2.7. Omnibus data fusion model

Bedworth and O'Brien [61] propose an alternative framework known as the Omnibus model. This hybrid model integrates elements from the Boyd loop, Dasarathy, and Waterfall models. As shown in Fig. 12, the Omnibus model comprises four primary modules, each designed to tackle different tasks in data fusion and achieve specific functional objectives. The Boyd control loop is described as an iterative process consisting of four stages: observe, orient, decide, and act, functioning in a continuous loop. In contrast, the Dasarathy model includes three fundamental levels of data fusion: data, feature, and decision levels. The Omnibus model combines these approaches to provide a comprehensive framework for data fusion.

3.2.8. Distributed data fusion using blackboard architecture

Schoess and Castore [62] present a distributed blackboard model for data fusion. As illustrated in Fig. 13, this model connects two sensors (s1 and s2) to multiple transducers (T). These sensors are overseen by a supervisor who manages conflicting sensor measurements, often based on the confidence levels attributed to each sensor. The transducers collect comprehensive information from the physical system under study, such as temperature and pressure. The fusion algorithm then generates a value, F, which relies on the data from the two sensors. Supervisors assign confidence levels to each sensor's readings. This approach can be conceptualized as a database containing sensory information, facilitating communication between various knowledge sources.

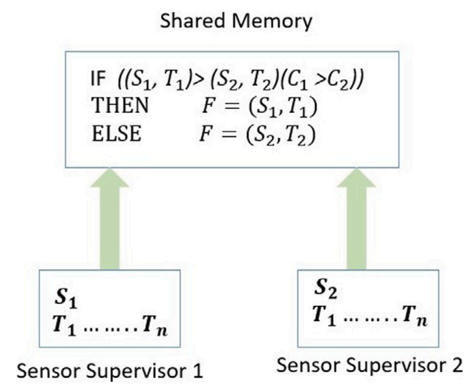

Fig. 13. The blackboard system for distributed data fusion.

Table 3
Symbols referenced in this section.

Symbol	Description
P(H)	Event probability H
m(·)	Degree of evidence support
$\mu_F(x)$	Function for fuzzy membership
$B^*(T)$	Approximation of T with conservatism
Θ	Framework for recognition
E	Available evidence
T	Set of interest in rough set theory
$B^*(T)$	Upper bound approximation of T

3.3. MSIF theories and methodologies

3.3.1. Traditional approaches in fusion theory

Classical information fusion theory primarily relies on mathematical techniques for statistical reasoning and estimation. These methods are particularly useful for processing and integrating incomplete data, such as inconsistent data types, low data credibility, and incomplete information. Table 3 presents a compilation of commonly utilized symbols within this specific section.

a. Fusion Through Probabilistic Modeling

Fusion methods based on probabilistic modeling were among the first theories applied in information fusion and remain the most widely used standard in the field [63]. These methods typically utilize Bayesian rules to combine prior knowledge with observational data. By employing probabilities to describe both the observed information and the necessary processing steps, these methods apply specific rules to synthesize this information into final decisions and comprehensive descriptions. In practice, probabilistic data fusion often hinges on Bayesian principles to merge prior information with new observations. Consequently, these methods are frequently referred to as "Bayesian fusion". However, there are multiple approaches to combining probabilistic information. Techniques such as Kalman filters, extended filters, sequential Monte Carlo methods, and density estimation functions can all serve as rules for multisource information fusion [64]. As a statistical inference method, Bayesian probability theory unifies all types of uncertainties into a single probability measure. It calculates the probability of a hypothesis being true by combining the prior probability of the hypothesis with the conditional probability of observed events. In smart sustainable cities, this can apply to scenarios like predicting low traffic congestion (H1) or high traffic congestion (H2), with relevant observations (evidence E) such as vehicle counts, weather conditions, and road incidents. Mathematically, Bayesian inference is represented as:

$$\begin{cases} P(H_i \mid E) = \dfrac{P(E \mid H_i)P(H_i)}{\sum_i P(E \mid H_i)P(H_i)} = \dfrac{P(E \mid H_i)P(H_i)}{P(E)} \\ P(H_1) + P(H_2) = 1, \text{ and } i \in [1,2] \end{cases} \tag{1}$$

where, $P(H_i \mid E)$ denotes the posterior probability that the hypothesis H_i is true given the evidence E (e.g., the probability of high traffic congestion after observing vehicle counts), $P(H_i)$ indicates the prior probability that the hypothesis H_i is true, and P(E) represents the conditional probability of observing the evidence E. When multiple pieces of evidence $E_1, E_2, \dots, E_n$ are present, the Bayesian estimator recursively combines them to update the probability distribution or density of the system's state or decision outcome [65]. For example, in determining whether air quality is good (H_1) or poor (H_2), multiple observations like PM2.5 levels,[1] temperature and humidity can be combined. The probabilities of these observations are denoted as $p(E_1)$, $p(E_2)$, …, $p(E_n)$. The probability of poor air quality can be calculated by integrating these observations.

$$P(H_1 \mid E_1, E_2, \dots, E_n) = \frac{P(H_1)\prod_{i=1}^{n} P(E_i \mid H_1)}{P(E_1, E_2, \dots, E_n)} \tag{2}$$

However, evaluating the prior distribution and normalization term often involves complex integrals, which limits the direct application of probability-based fusion methods. The Kalman filter, a type of Bayesian filter, simplifies this by providing an exact analytical solution, making the computation of fusion results easier. This simplicity and ease of implementation have made the Kalman filter a popular choice for probabilistic fusion. However, like other least squares estimators, the Kalman filter is sensitive to outliers, making it unsuitable for applications with poorly parameterized error characteristics. Therefore, extensions of the Kalman filter, which often use approximation techniques, are employed to handle data fusion in nonlinear systems.

For example, the Extended Kalman Filter (EKF) [66] and the Unscented Kalman Filter (UKF) [67] address data processing challenges in nonlinear systems by applying first- and second-order approximations to the original Kalman filter. However, these methods can only manage nonlinear problems within a limited scope due to their approximation range. To address this, Stone et al. [68] proposed a grid-based method to approximate nonlinear probability density functions. Nonetheless, this approach faces a rapid increase in computational complexity as data dimensionality rises. To address nonlinear system data fusion problems, extensions of the Kalman filter, such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF), apply first- and second-order approximations. However, these methods are limited to handling nonlinear problems within a certain scope. Stone et al. proposed a grid-based method to approximate nonlinear probability density functions, but this approach suffers from increased computational complexity with higher data dimensionality.

To mitigate the "curse of dimensionality", the Markov Chain Monte Carlo (MCMC) algorithm [69] was developed. This algorithm reduces the complexity of approximating high-dimensional densities by evolving samples through a Markov chain, rather than relying on independent random computations at each step. In this framework, the Markov chain consists of a sequence of random samples generated according to a transition probability (kernel) function with Markov properties. This means that the transition probabilities between sample values rely solely on the current state of the sample. Consequently, well-constructed Markov chains can converge to a specific stationary distribution with respect to the sampled values.

b. Theory of Belief Functions

Belief functions theory, rooted in Dempster's work [70] on the reliability of source states in multisource information fusion (MSIF) and formalized by Shafer, provides a framework for evidence-based reasoning [71]. This theory encompasses Dempster–Shafer Theory [72] (DST) and Dezert–Smarandache Theory [73] (DSmT), which assign beliefs and plausibilities to measurement hypotheses and define rules for fusing uncertain and imprecise data. These concepts extend Bayesian theory by handling probability mass functions. Garvey et al. [74] first applied this theory to information fusion in 1981.

In MSIF [75], the term 'beliefs' includes indicators like reliability, truthfulness, and dependence:

-*Reliability* in MSIF is the consistency and stability of information from a source. Reliable sources provide consistent, interference-resistant data, while unreliable ones can lead to inaccurate judgments and predictions, affecting system accuracy and credibility.

-*Truthfulness* measures how closely information aligns with reality. False data can deceive the system, resulting in unreliable outcomes and significantly impacting performance.

-*Dependence* in MSIF refers to the interconnectedness of sources, where one source's information influences another. Properly accounting for these interdependencies can enhance the system's robustness and stability, while overreliance on a single source can lead to imbalanced and inaccurate results.

In practical use, unlike probability fusion based on Bayesian inference, belief functions theory has the advantage of providing information at different levels of detail. For example, in tasks involving human behavior recognition, the results can range from broad categories such as static and motion states to specific actions like lying, sitting, standing, walking, running, and jumping [76]. To illustrate, consider a set $\Theta = \{\theta_1, \theta_2, \dots, \theta_6\}$ representing these six postures. Using a basic belief assignment function m (Adhering to Eq. (3)), confidence levels are systematically assigned to all potential recognition results:

$$m(\emptyset) = 0, \sum_{H \subseteq 2^{\Theta}} m(H) = 1 \tag{3}$$

Here, $2^{\Theta} = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \dots, \{\theta_6\}, \{\theta_1 \cup \theta_2\}, \{\theta_1 \cup \theta_3\}, \dots, \Theta\}$ denotes the power set of Θ, which includes all subsets. The function $m(H)$ indicates the degree of support for different propositions based on current evidence E. Impractical states, such as "both running and lying", have belief values set to zero, while valid states are referred to as focal elements.

When combining multiple sources of information, each source is treated as independent evidence with its own basic belief assignment function. For example, if several sensors collect data simultaneously, each sensor's output is an independent evidence source E_i with a corresponding belief assignment function m_i. The integration and processing of multiple information sources can be accomplished using Dempster's rule for evidence combination [77], as described in Eq. (4). In this context, K denotes the conflict coefficient, which quantifies the degree of disagreement among recognition results from different evidence sources. A higher K value indicates greater conflict. If K equals 1, the rule becomes ineffective.

$$\begin{cases} m(H) = \dfrac{\sum_{H_i} \prod_{i=1}^{n} m_i(H_i)}{1-K}, \quad 1 \le i \le n \text{ and } H \neq \emptyset \\ m(\emptyset) = 0, \\ K = \displaystyle\sum_{H=\emptyset} \prod_{1 \le i \le n} m_i(H_i) \end{cases} \tag{4}$$

In this Eq. (4), $m(H)$ represents the combined belief assignment for proposition H. The numerator calculates the belief by considering all possible subsets H_i of H and combining the belief assignments from each evidence source. Dempster's rule provides a structured method for integrating multiple evidence sources, resolving conflicts, and forming a unified belief assignment.

While Dempster's rule in Dempster–Shafer Theory (DST) is effective for managing uncertain information, it can produce counterintuitive

[1] PM2.5 refers to particulate matter with a diameter of 2.5 micrometers or smaller, often used as an indicator of air pollution levels.

results when dealing with highly conflicting evidence in multisource information fusion (MSIF). To address this, Xiao [78] proposed the Generalized Evidential Jensen–Shannon (GEJS) divergence to measure conflicts among multiple evidence sources. This method enhances precision by assigning appropriate weights to each source, resulting in a weighted average evidence that is fused using Dempster's Combination Rule (DCR) to support robust decision-making.

In comparison, Dezert–Smarandache Theory (DSmT) excels at handling non-exclusive elements, similar to D-Numbers Theory (DNT) [79] and order-2 fuzzy sets [80]. DNT expands on DST by addressing limitations related to non-exclusive elements. It involves combining evidence from various sources into basic belief assignments using D-functions [81], which are then processed through combination, normalization, and uncertainty reassignment steps. DNT is advantageous in dealing with complex uncertainties and non-exclusive elements but requires significant computational effort and sufficient data.

Additionally, Complex Evidence Theory (CET) has gained attention for its interpretability in MSIF applications. CET introduces the Complex Evidence Correlation Coefficient (CECC) for measuring conflicts among Complex Basic Belief Assignments (CBBA), leading to the development of algorithms like CECC-WDMSIF for improved expert system performance. CET also incorporates quantum frameworks to enhance uncertainty reasoning and knowledge representation. Moreover, the interpretability of Complex Evidence Theory (CET) has attracted attention within the realm of MSIF applications. For instance, in a recent study [82] (Ref. [72]), researchers focused on enhancing the performance of expert systems by introducing a novel metric called the Complex Evidence Correlation Coefficient (CECC). This coefficient enables the measurement of conflicts among Complex Basic Belief Assignments (CBBA) within CET. Additionally, a weighted discounting multisource information fusion algorithm, CECC-WDMSIF, was devised based on CECC to bolster the capabilities of CET-based expert systems. Furthermore, another study [83] explored the integration of quantum frameworks into CET to address uncertainty in knowledge representation. This endeavor led to the proposal of a generalized negation method for quantum basic belief assignments, offering insights into effective uncertainty measurement and quantum information fusion. Unlike probabilistic modeling, which assumes uniform distributions, belief functions theory assigns probabilities only when there is supporting information. This allows for expressing confidence in uncertain information by allocating the entire mass to the discrimination framework. Therefore, choosing between probabilistic modeling and belief functions theory depends on balancing the need for data accuracy with the flexibility of fusion formulas (The theory of belief functions) [84].

c. Theory of Fuzzy Sets

Fuzzy set theory is an effective framework for managing and integrating data in multisource information fusion (MSIF) tasks, especially when dealing with complex and uncertain information. Unlike probability-based methods, fuzzy set theory excels in handling vague and imprecise data, making it particularly useful when the definitions of objects or situations are not clear-cut [85]. In the context of MSIF for smart sustainable cities, this capability is invaluable for dealing with the diverse and often ambiguous data streams generated by various urban sensors and systems.

A first-order membership function $\mu_A(x)$ represents the degree to which an element x belongs to a fuzzy set A, with values ranging from 0 to 1 [85]. This concept is extended by second-order membership functions, which describe the distribution of the first-order membership function across the entire membership space [86].

Operations in fuzzy set theory [87]., such as conjunction (AND) and disjunction (OR), enable the precise combination of fuzzy sets. For conjunction, the membership degree of x in the conjunctive set $A \cap B$ is the minimum of its membership degrees in the two fuzzy sets:

$$\{\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x)) \tag{5}$$

For disjunction, the membership degree of x in the disjunctive set $A \cup B$ is the maximum of its membership degrees in the two fuzzy sets:

$$\{\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x)) \tag{6}$$

Adaptive fuzzy fusion rules offer a versatile method for managing data from conflicting or unreliable sources, which is particularly useful in scenarios with high levels of conflict. This approach is vital for information fusion within the context of smart sustainable cities, where diverse data types from traffic sensors, air quality monitoring systems, and social media need to be synthesized to ensure thorough situational awareness [88]. In the application of Multisource Information Fusion (MSIF) for smart sustainable cities, the fuzzy set theory presents several benefits:

– **Handling Imprecise Data**: This theory is adept at modeling and integrating data that may be uncertain or imprecise, often found in human-generated inputs and subjective evaluations [85]. For instance, information related to traffic congestion or energy consumption often lacks precision.
– **Linguistic Data Integration**: It facilitates the fusion of linguistic descriptions like "very high", "medium", or "low", which is advantageous for integrating expert insights and subjective decisions into urban planning [89].
– **Enhanced Uncertainty Handling**: Utilizing second-order membership functions, this theory allows for a more precise description of the relationships among different data sets, enhancing the handling and fusion of uncertainties [86]. This capability is essential when dealing with data variability in environmental sensors, public health records, and other urban management systems.

To further illustrate the application of fuzzy set theory, consider the following equations for a second-order membership function $\mu_B(\mu_A(x))$:

$$\{\mu_B(\mu_A(x)) = \mu_B(y), \quad \text{where} \quad y = \mu_A(x) \tag{7}$$

Additionally, the representation of adaptive fuzzy fusion rules can be expressed as:

$$\{\mu_C(x) = \alpha\mu_A(x) + (1-\alpha)\mu_B(x), \quad 0 \le \alpha \le 1 \tag{8}$$

In the realm of smart sustainable cities, fuzzy set theory assists in handling specific challenges such as [88]:

– **Traffic Management**: By integrating data from various traffic sensors and public transportation systems, fuzzy set theory can help optimize traffic flow and reduce congestion through adaptive traffic signal control and real-time route planning.
– **Environmental Monitoring**: Fuzzy set theory aids in fusing data from multiple environmental sensors to provide accurate and timely information about air and water quality, enabling proactive measures to address pollution.
– **Energy Management**: It supports the integration of data from smart grids, renewable energy sources, and consumer usage patterns to optimize energy distribution and reduce wastage.

Overall, the fuzzy set theory provides a robust method for managing and integrating uncertain and fuzzy data, making it a valuable tool for addressing the challenges posed by ambiguous object definitions and incorporating human expertise in MSIF tasks, particularly in the context of smart sustainable cities.

d. Theory of Rough Set

Rough set theory is an effective mathematical approach for handling imprecise, uncertain, and incomplete data. This makes it particularly valuable for Multi-Source Information Fusion (MSIF) in smart sustainable cities. Unlike probabilistic models, which require prior knowledge, rough set theory works directly with the available data. It enables the analysis of data granularity and effectively manages uncertainty at various levels of detail, facilitating robust decision-making in complex urban environments [90]. In rough set theory, the approximation of a target set T within a given framework F_B for a set $B \subseteq A$ is expressed as $\langle B^*(T), B_*(T)\rangle$. Here, $B^*(T)$ and $B_*(T)$ represent the upper and lower approximations of the set T, respectively, where F_B describes a specific set of attributes for the objects. The lower approximation $B_*(T)$ includes all elements that definitively belong to T, while the upper approximation $B^*(T)$ includes all elements that possibly belong to T. These can be mathematically represented as:

$$\begin{cases} B_*(T) = \{x \in U \mid [x]_B \subseteq T\} \\ B^*(T) = \{x \in U \mid [x]_B \cap T \neq \emptyset\} \end{cases} \tag{9}$$

where U is the universe of discourse and $[x]_B$ is the equivalence class of x under the indiscernibility relation B [90]. Additionally, the boundary region, which represents the elements that cannot be decisively classified as belonging or not belonging to T, is given by:

$$\left\{ B_{BN}(T) = B^*(T) - B_*(T) \right. \tag{10}$$

In smart sustainable cities, rough set theory aids in efficiently handling the complexities and uncertainties inherent in data from diverse sources. For instance, in urban traffic management, rough set theory helps approximate traffic patterns, thereby managing uncertainties related to traffic flow, congestion, and incidents. This allows city planners to make more informed decisions based on incomplete or uncertain data, ultimately improving traffic efficiency and reducing congestion [91].

Environmental monitoring is another critical application of rough set theory in smart cities. By approximating pollution levels and managing uncertainties in sensor data collected from air quality monitors, noise sensors, and weather stations, rough set theory provides a robust assessment of environmental conditions. This enables city planners to implement more effective pollution control measures and improve urban living conditions [92]. Moreover, rough set theory can enhance the integration and interpretation of data from smart grids, renewable energy sources, and consumer usage patterns, optimizing energy distribution and reducing wastage. This integration supports the development of sustainable energy management practices in smart cities, ensuring a more reliable and efficient energy supply [93]. To quantify the dependency between attributes in rough set theory, the dependency degree γ is used, which measures how much the decision attribute depends on a set of condition attributes. It is defined as:

$$\left\{ \gamma_B(T) = \frac{|B_*(T)|}{|U|} \right. \tag{11}$$

where $|B_*(T)|$ is the cardinality of the lower approximation set and $|U|$ is the total number of objects in the universe. Furthermore, the discernibility matrix is a useful tool in rough set theory for feature selection and data reduction. It is defined as follows for a dataset with objects x_i and x_j:

$$\left\{ DM(x_i, x_j) = \{a \in A \mid a(x_i) \neq a(x_j)\} \right. \tag{12}$$

where $DM(x_i, x_j)$ represents the set of attributes that can distinguish between objects x_i and x_j.

Overall, the rough set theory provides a comprehensive framework for dealing with the intricacies of multisource information fusion in smart sustainable cities, providing robust solutions for traffic management, environmental monitoring, and energy management. Its ability to handle uncertain and incomplete data makes it an invaluable tool for urban planners and policymakers aiming to create more efficient and sustainable urban environments.

e. Affinity Propagation Clustering for Multisource Information Fusion

Affinity Propagation (AP) is an effective clustering algorithm [94] for multisource information fusion (MSIF) in smart sustainable cities. Unlike traditional methods such as k-means, which require predefined cluster numbers, AP dynamically determines the number of clusters based on the data itself. This capability is particularly beneficial in MSIF applications, where the exact number of underlying data sources may be unknown, allowing for more flexible and accurate data integration and analysis.

- How Affinity Propagation Works

Affinity Propagation (AP) treats each data point as a potential exemplar and iteratively exchanges messages to determine the best set of exemplars and corresponding clusters. The algorithm utilizes two types of messages:

1. **Responsibility (r(i, k))**: Sent from data point i to candidate exemplar k, indicating the suitability of k as an exemplar for i.
2. **Availability (a(i, k))**: Sent from candidate exemplar k to data point i, indicating the appropriateness of i choosing k as its exemplar.

The message updates are computed as follows:

– **Responsibility update**:

$$r(i,k) \leftarrow s(i,k) - \max_{k' \neq k}\{a(i,k') + s(i,k')\} \tag{13}$$

– **Availability update**:

$$a(i,k) \leftarrow \min\{0, r(k,k) + \sum_{i' \notin \{i,k\}} \max\{0, r(i',k)\}\} \tag{14}$$

- Advantages of Affinity Propagation

– **Automatic Cluster Number Determination**: AP does not require a predefined number of clusters, which is advantageous for MSIF where the exact number of data sources or clusters is unknown in advance.
– **Compatibility with Non-Euclidean Distances**: AP can operate with various similarity measures, making it adaptable for different types of data.
– **Scalability**: The algorithm is computationally efficient and capable of handling large datasets, making it ideal for real-time applications in smart cities.

- Application in Smart Sustainable Cities

In smart sustainable cities, AP can be applied to cluster data from various sensors and information sources. For instance, it can integrate and cluster data from traffic sensors, pollution monitors, weather stations, and social media feeds to identify patterns and anomalies. This clustering can help city planners optimize traffic flow, control pollution, and manage resources effectively. For example, consider a smart city where data is collected from multiple sensors monitoring air quality, traffic flow, and weather conditions. Using AP, the similarity $s(i,k)$ between two data points (e.g., measurements from two different sensors) could be based on a combination of factors such as the type of sensor, the geographical distance between sensors, and the time of measurement. The responsibility and availability messages would then iteratively update to identify clusters of sensors that provide similar information, helping to detect areas with high pollution levels or traffic congestion. By leveraging Affinity Propagation, smart city planners can gain deeper insights from the integrated data, leading to more informed decision-making and improved urban management. This approach enhances the robustness and accuracy of MSIF, facilitating the development of more efficient and sustainable urban environments.

f. Theory of Possibility for Multisource Information Fusion

Possibility Theory is an extension of fuzzy set theory that deals with the concept of possibility distributions, providing a measure of how possible different outcomes are. It is particularly useful in situations where information is incomplete or uncertain, making it a valuable tool for MSIF in smart sustainable cities.

- How Possibility Theory Works

Possibility theory employs two main functions to handle uncertainty: the possibility measure Π and the necessity measure N. These functions are used to evaluate the degree of possibility and necessity of events, respectively.

– **Possibility Measure** Π: The possibility measure of an event A is defined as:

$$\left\{ \Pi(A) = \sup_{x \in A} \pi(x) \right. \tag{15}$$

where $\pi(x)$ is the possibility distribution, indicating the degree of possibility of each outcome x.

– **Necessity Measure** N: The necessity measure of an event A is defined as:

$$\left\{ N(A) = 1 - \Pi(A^c) \right. \tag{16}$$

where A^c is the complement of A. The necessity measure indicates the degree to which A is necessarily true.

- *Advantages of Possibility Theory*

– **Handling Incomplete Information**: Possibility theory is effective in managing situations where data is incomplete or missing, providing robust measures of uncertainty.
– **Flexibility**: The theory allows for the representation of both complete certainty and complete ignorance, offering a flexible framework for dealing with various types of uncertainty.
– **Integration with Fuzzy Set Theory**: Possibility theory extends fuzzy set theory, making it well-suited for applications involving fuzzy data and linguistic variables.

-*Application in Smart Sustainable Cities*

In the context of smart sustainable cities, possibility theory can be used to integrate and analyze data from various sources with inherent uncertainty. For instance, it can be applied to environmental monitoring, where data from sensors measuring air quality, noise levels, and weather conditions may be incomplete or uncertain. Possibility theory can help to quantify the uncertainty and provide meaningful insights for city planners.

For example, consider a smart city scenario where data from air quality sensors is used to monitor pollution levels. The possibility distribution $\pi(x)$ could represent the degree of possibility of different pollution levels based on sensor readings. City planners can use the possibility and necessity measures to assess the likelihood and necessity of high pollution levels and take appropriate actions to mitigate the impact. By leveraging possibility theory, smart cities can improve their ability to handle uncertain and incomplete data, leading to better decision-making and more effective urban management. This approach enhances the capability of MSIF to provide reliable and comprehensive insights in complex urban environments.

-In addition to the established MSIF methodologies, recent studies have introduced innovative fusion techniques that significantly enhance the management of uncertainty and conflicting evidence within MSIF frameworks. Notably, the development of a novel Quantum Dempster's Rule of Combination (QDRC) [95] integrates quantum computing principles into Dempster–Shafer evidence theory. This approach addresses the exponential increase in computational complexity typical of traditional Dempster's rule as the number of elements in the identification framework grows. By utilizing quantum circuits, such as the Toffoli gate, QDRC significantly reduces this complexity, enabling more efficient and accurate pattern classification, especially in real-time and large-scale applications. The quantum-enhanced method is demonstrated through experiments on quantum cloud platforms like IBM and IonQ, showcasing improved accuracy as the number of measurements increases.

Furthermore, the introduction of the Fractal Belief Jensen–Shannon (FBJS) divergence [96] and Higher Order Belief Jensen–Shannon (HOBJS) divergence [97] provide refined mechanisms for handling conflicting evidence, enhancing the accuracy of multisource information fusion processes. These methods are particularly effective in environments where data sources frequently conflict, offering robust solutions for resolving these disputes and improving pattern classification outcomes. Lastly, the Higher Order Fractal Belief Rényi Divergence (HOFBReD) [97] combines the principles of fractal structures and Rényi divergence to offer a more detailed and accurate quantification of belief discrepancies, proving its value in complex and uncertain environments.

3.4. Algorithms, data structures, and computational complexity

This subsection delves into the critical algorithms, data structures, and computational complexities that underpin multisource information fusion (MSIF) techniques, providing a comprehensive examination of these elements in the context of smart sustainable cities.

3.4.1. Algorithms employed in MSIF

a. Data-Level Fusion:

- *Convolutional Neural Networks (CNNs):* CNNs are extensively utilized in image data fusion, particularly in urban surveillance and environmental monitoring. These networks leverage spatial hierarchies in image data, enabling the effective combination of inputs from multiple sources such as camera feeds and satellite imagery. The convolutional layers within CNNs apply filters to local regions of the input data, facilitating efficient processing of large-scale image data—an essential requirement for real-time applications where speed and accuracy are paramount.

- *Kalman Filters:* Kalman filters play a crucial role in dynamic systems requiring continuous data fusion, such as real-time traffic monitoring or sensor networks within smart cities. These filters offer a recursive solution to the linear quadratic estimation problem, allowing for the prediction of a system's state and its continuous update as new data becomes available. The computational complexity of Kalman filters is linear with respect to the number of state variables, making them particularly suitable for applications demanding quick and accurate updates in real-time environments.

b. Feature-Level Fusion:

- *Principal Component Analysis (PCA):* PCA is a statistical technique employed to reduce the dimensionality of large datasets while retaining the most significant variance. This method is particularly beneficial in urban data analytics, where diverse data from sources like air quality sensors and traffic monitors need to be combined efficiently. PCA simplifies the fusion process by identifying the most critical features in the data, thereby reducing computational complexity and enabling the processing of large datasets without overwhelming computational resources.

- *Autoencoders:* Autoencoders are neural networks designed to learn efficient codings of data. In the context of smart cities, autoencoders compress large datasets, such as sensor readings, into lower-dimensional representations before fusion. This compression reduces the computational load during the fusion process, particularly when dealing with high-dimensional data, making it an effective tool for managing large volumes of information in urban environments.

c. Decision-Level Fusion:

- *Ensemble Learning Methods:* Ensemble learning methods, including Random Forests and Gradient Boosting Machines, combine the outputs of multiple classifiers to improve decision-making accuracy. In smart city applications – where decisions must be made based on a wide array of data inputs, such as security measures or traffic management – these methods ensure robustness and reliability. The trade-offs between the accuracy of these methods and their computational requirements will be examined, especially in scenarios involving large datasets and numerous classifiers.

3.4.2. Data structures

- Hierarchical Data Structures:

Tree-Based Structures: Tree-based data structures are essential for efficiently managing and retrieving large datasets within MSIF systems. For example, hierarchical clustering in environmental data analysis benefits from tree-based structures, which organize data points into nested clusters. This organization enables quick access and processing of data, making it particularly useful in applications like urban planning and resource management, where large volumes of diverse data must be handled efficiently.

- Graph-Based Structures:

Graphs: Graph-based structures are powerful tools for representing complex relationships between data sources and fusion levels. In the networks of a smart city, these structures model the interactions between different entities, such as sensors and databases, and facilitate efficient data fusion across multiple layers. Graph algorithms, such as Dijkstra's or A* for shortest path finding, are applied to optimize data flow and processing within MSIF systems, ensuring that the fusion process is both accurate and scalable.

3.4.3. Computational complexity

- Complexity Analysis

- *Time and Space Complexities:* A detailed analysis of the time and space complexities associated with the algorithms and data structures discussed will be provided. For instance, the complexity of CNNs will be examined concerning their depth and the size of their input tensors, while the complexity of Kalman filters will be analyzed in relation to the number of state variables and the speed of convergence. This analysis will also address the challenges of scaling these methods to handle the large volumes of data typical in smart city applications.

-*Scalability Challenges:* As smart cities generate massive amounts of data at high velocities, ensuring that MSIF methods can scale effectively is critical. The computational challenges posed by this scale, such as memory constraints and processing time, will be discussed, along with potential mitigations through algorithmic optimizations and efficient data management practices.

- Optimization Strategies:

- *Algorithmic Refinements:* To reduce computational load, refinements such as pruning techniques in decision trees and approximations in Bayesian networks will be explored, which maintain accuracy while lowering computational costs.

- *Hardware Accelerations:* The use of GPUs and specialized hardware for accelerating deep learning models and other computationally intensive tasks will be discussed, particularly in the context of real-time applications where processing speed is crucial.

- *Distributed Computing Frameworks:* The application of cloud, fog, and edge computing frameworks will be detailed, highlighting how these frameworks can distribute computational tasks across multiple nodes, thereby enhancing the efficiency and responsiveness of MSIF systems in smart cities.

4. Conceptual foundations for MSIF in smart sustainable cities

Multisource Information Fusion (MSIF) is crucial in urban settings due to its ability to amalgamate extensive and varied data from multiple sources, such as environmental sensors, public records, traffic management systems, and citizen feedback. This fusion facilitates a more integrated view of urban life, enabling city authorities and planners to make informed decisions that enhance urban operations, improve resource efficiency, and elevate the overall quality of life in smart sustainable cities. Effective MSIF is instrumental in addressing urban issues like congestion, safety, environmental sustainability, and energy management, contributing to the development of cities that are both adaptive and resilient.

The dynamic and complex nature of urban data necessitates a robust MSIF framework capable of handling these challenges adeptly. To this end, we introduce the Advanced Hierarchical and Multi-Dimensional Classification Framework, specifically tailored for modern urban settings.

4.1. Our advanced hierarchical and multi-dimensional classification framework

Our framework is strategically developed to manage the intricate and evolving demands of urban environments through sophisticated MSIF processes. It organizes data fusion into multiple dimensions, each designed to enhance specific aspects of urban management:

a. Fusion Objectives (O)

- **O1: Resolving Problematic Data**
  Problematic data issues [98,99] such as inconsistency and imperfection are addressed through data fusion, enhancing overall data quality [100,101]. This approach is crucial in managing the complex data environments typical in smart cities.
- **O2: Enhancing Data Reliability**
  Data fusion improves reliability [102,103], particularly in less controlled environments with high noise levels. Adding redundancy [104] through multiple data sources is a key strategy for enhancing data reliability and security in smart cities.
- **O3: Deriving Higher Level Insights**
  Advanced data mining and fusion techniques are employed to extract actionable insights from complex urban datasets [1,105]. These techniques are pivotal for strategic urban planning and decision-making.
- **O4: Improvement of data completeness**
  Data fusion is used to overcome coverage limitations by integrating multiple data sources [106], ensuring a comprehensive view of urban dynamics [107].
- **O5: Optimizing Resource Allocation**
  By integrating and analyzing data [108] from various sources, smart sustainable cities can optimize resource distribution [109] and utilization. This strategic, data-driven decision-making enhances operational efficiency and sustainability across urban sectors.
- **O6: Enhancing Predictive Capabilities**
  The fusion of multisource data significantly boosts the predictive analytics [110] capabilities of smart cities [1]. This enhancement allows for more accurate forecasting of urban dynamics like traffic flows and population movements, facilitating proactive urban planning and response strategies.

b. Fusion Techniques (T)

- **T1: Associating Data**
  Data association is a vital fusion technique that consolidates information from various sources based on similarities, improving the accuracy and reliability of insights. By combining correlation analysis with machine learning models [111] like Support Vector Machines (SVM), significant relationships among datasets are uncovered, enhancing their utility for decision-making. Techniques such as Nearest Neighbors [112], Probabilistic Data Association [113], and Multiple Hypothesis Testing [114] ensure a comprehensive understanding of the data landscape, leading to precise and effective outcomes.

T2: Estimating States
State estimation uses multiple data sources to enhance the accuracy of system state predictions. Techniques like Kalman Filters [115] and Hidden Markov Models (HMM) [116] are employed for this purpose. For example, the equation $x^{t+1} = Ax^t +$

$Bu^t + w$, where w represents process noise, illustrates their application. Common state estimation techniques include Particle Filter [117], Kalman Filter [115], Maximum Likelihood [118], and Covariance Consistency Model [119]. These methods collectively improve the precision and reliability of state predictions, crucial for applications such as navigation systems and financial modeling.

- **T3: Integrating Decisions**
  Decision fusion involves integrating decisions made by various sub-components of a system to achieve a unified objective. For instance, in a smart sustainable city, decision fusion can be used to optimize energy management [120] by combining decisions from different modules, such as renewable energy sources, grid management systems, and consumption monitoring [121]. This process leverages decision-making algorithms, such as ensemble methods and Bayesian networks, to enhance decision-making efficiency across urban management platforms. Common techniques in decision fusion include Dempster–Shafer Inference [122], semantic approaches [123], and Bayesian inference [124]. These methods collectively improve decision accuracy and reliability by synthesizing information from multiple sources and perspectives.
- **T4: Classifying and Predicting**
  Advanced machine learning techniques, particularly deep learning networks [125], are applied to project urban [105] and categorize data, facilitating proactive management.
- **T5: Dimensional Reduction Techniques**
  Dimensional reduction aims to decrease the number of variables in a dataset while preserving essential information. Techniques such as T-distributed Stochastic Neighbor Embedding [126] and Principal Component Analysis [127] are employed for feature extraction and visualization. These methods reduce computational complexity and help manage high-dimensional data challenges, thereby improving the efficiency and interpretability of data analysis.
- **T6: Statistical Techniques for Inference**
  Statistical inference [128] utilizes probabilistic methods to draw insights from data. Techniques like regression analysis, hypothesis testing [129], and Bayesian inference are employed to predict outcomes and uncover relationships. These methods are essential for data-driven decision-making, as they allow for the estimation of population parameters and the validation of theoretical models using observed data.
- **T7: Data Representation**
  Data representation [130] involves converting complex datasets into visual formats for better comprehension. Techniques such as scatter plots, heatmaps, and interactive dashboards facilitate the identification of patterns, trends, and anomalies. Effective visualization [131] enhances the accessibility and usability of data, aiding in informed decision-making processes.

c. Data Types and Outputs (D)

Dasarathy's classification system [132] organizes data fusion methods by the relationship between input and output data. It includes five distinct categories, as shown in Table 4:

Each class addresses specific data fusion requirements, from basic data processing to complex decision-making, ensuring a comprehensive understanding and application of these techniques.

d. Data Sources (S)

In sustainable city applications, data sources can be categorized into four main types, regardless of their communication medium:

- **S1: Sensor-Based Data Sources**
  These originate from sensors [1] deployed to monitor specific locations or city-wide areas. Examples include traffic sensors, noise level meters, weather stations, and smart meters for utilities like water and electricity. For instance, a smart meter [133] that tracks electricity usage and sends data to a central system is considered a physical data source.
- **S2: Cyber-Based Data Sources**
  These datasets are primarily obtained from the internet [134], such as e-commerce transaction records, online forums [135], and digital service usage logs. Data mining techniques are commonly used to gather this information from APIs or third-party vendors. Examples include data from online shopping behavior, internet traffic logs, and streaming service usage patterns.
- **S3: User-Contributed Data Sources**
  Collected from individuals via personal devices like mobile phones [136], fitness trackers, and smart home devices, this data is contributed voluntarily or through incentive programs [137]. Examples include health data from fitness apps, location data from mobile GPS, and user feedback on municipal services collected through mobile apps.
- **S4: Hybrid and Integrated Data Sources**
  Combining data from multiple sources, such as integrating participatory and physical sensor data, these sources provide more comprehensive insights [138]. Examples include combining data from environmental sensors with user-reported air quality perceptions, or merging traffic sensor data with crowdsourced traffic reports from navigation apps [139]. Hybrid data sources leverage the strengths of various data types to enhance analysis and decision-making in smart city environments.

Each type of data source plays a critical role in the multifaceted data ecosystem of smart sustainable cities, contributing to a more complete and nuanced understanding of urban dynamics.

e. Fusion Scales (L)

Data fusion scales are categorized based on sensor coverage rather than deployment, crucial for understanding integration levels in smart sustainable cities. These scales include:

- **L1: At the Sensor Level**
  At this level, data from individual sensors are processed directly at the source using edge computing solutions to reduce latency. An example is the fusion of data from various smartphone sensors [140].
- **L2: Building Wide Fusion**
  Data collected within a building [141] is integrated to form comprehensive outputs. For instance, combining building energy data with security systems to create a building management system.
- **L3: Inter-Building Fusion**
  Data from multiple buildings within a small area [142], such as a university campus, are combined to generate specific outputs. This scale of fusion helps in managing multiple premises collectively.
- **L4: City Wide Fusion**
  Extensive data fusion across an entire city involves both centralized and decentralized computing architectures. This integration supports city-wide studies, such as analyzing citizen behavior based on data collected from different city areas [143].
- **L5: Inter-City Fusion (or Larger)**
  This scale involves fusing data from large areas, including multiple cities or diverse terrains like mountains and forests. Examples include [144] comparing data between different smart cities or integrating data from city outskirts with surrounding regions.

f. Platform Architectures (P)

The architecture of computational platforms for data fusion is another crucial classification aspect. We identify four main types:

- **P1: Edge Computing**
  Data processing occurs close to the data source, using devices like microcontrollers and Raspberry Pi [145]. This approach reduces communication overhead and latency.

Table 4
Types of data and their outputs.

Category	Description	Techniques	Application
D1: Data In-Data Out (DAI-DAO)	Processes raw data as both input and output.	Signal and image processing algorithms.	Immediate post-sensor data collection, providing reliable results.
D2: Data In-Feature Out (DAI-FEO)	Extracts features or characteristics from raw data.	Feature extraction algorithms.	Describes environmental entities.
D3: Feature In-Feature Out (FEI-FEO)	Uses features as input and output to enhance or create new features.	Feature fusion, symbolic fusion.	Improves or generates features from existing data.
D4: Feature In-Decision Out (FEI-DEO)	Converts features into decisions.	Classification systems, decision algorithms.	Common in systems making decisions based on sensor inputs.
D5: Decision In-Decision Out (DEI-DEO)	Fuses multiple decisions to create improved outcomes.	Decision fusion algorithms.	Enhances decision-making by combining various inputs.

- **P2: Fog Computing**
  Data is processed at an intermediary layer between the edge and cloud [146]. This method offloads computation, suitable when stable power sources at the edge are unavailable [147].
- **P3: Cloud Computing**
  Data is processed and fused in the cloud, ideal for handling big data [148] with the advantages of easy data access. However, it incurs higher communication overhead and costs [149].
- **P4: Hybrid Computing**
  Combines edge, fog, and cloud layers [150]. Initial data processing occurs at the edge or fog, while higher-level analysis is conducted in the cloud, balancing resource use and application needs [151].

g. Adaptive Techniques (A) Adaptive techniques [152] in smart sustainable systems promote the ongoing enhancement and peak performance through self-adjustment and learning abilities [153]. These methods ensure that systems remain efficient and responsive to changing conditions.

- **A1: Dynamic Configurations**
  Systems self-adjust configurations in response to environmental changes or data variations, optimizing performance and efficiency through real-time adjustments [154].
- **A2: Continuous Learning and Evolution**
  Incorporates continuous learning to adapt and improve based on new data and past experiences, leveraging machine learning and AI to autonomously identify patterns and make adjustments [155].
- **A3: Predictive Adaptation**
  Utilizes predictive analytics to anticipate changes or disruptions, allowing proactive operational adjustments to maintain stability and efficiency [156].
- **A4: Self-Optimization**
  Employs algorithms for continuous optimization [157], ensuring peak performance without manual intervention.
- **A5: Context-Aware Adjustments**
  Uses context-awareness to tailor responses based on situational parameters, enhancing overall performance and user experience [158].
- **A6: Fault Tolerance and Recovery**
  Includes mechanisms for detecting, managing, and recovering from errors autonomously, ensuring uninterrupted and reliable operation [159].

h. Integration Complexity (I) The complexity of integration in multisource information fusion (MSIF) systems involves handling the challenges of merging diverse data sources and ensuring interoperability across different formats and systems. Tackling these complexities is crucial for the efficient functioning of smart sustainable cities.

- **I1: Managing Complexity**
  Utilizes advanced algorithms and data models to integrate various data sources smoothly. Techniques such as multi-agent systems and modular architectures help manage and streamline these complex integrations [160].
- **I2: Ensuring Interoperability**
  Achieves seamless interaction between different data formats and systems through [161] standardized protocols and middleware solutions like API gateways and data translation layers.
- **I3: Scalability**
  Ensures the integration framework can scale to accommodate increasing data volumes and new sources by using cloud-based infrastructures and scalable architectures [162].
- **I4: Data Security and Privacy Protection**
  Implements comprehensive security protocols and privacy-preserving methods to safeguard sensitive information during the integration process. This includes techniques such as encryption, data anonymization, and strict access control measures to ensure data integrity and confidentiality [163].

5. Challenges in multisource information fusion (MSIF) for smart sustainable cities

Implementing Multisource Information Fusion (MSIF) in smart sustainable cities involves addressing various technical challenges. These challenges arise from integrating diverse data sources, handling large volumes of data in real time, and ensuring reliable decision-making processes. Key areas of concern include data completeness, data heterogeneity, and the dynamic nature of urban environments. By understanding these challenges, researchers and practitioners can develop more effective MSIF systems to enhance urban sustainability and efficiency.

5.1. Current challenges in multisource information fusion: Data challenges in MSIF

a. Incomplete Data: Data from sensors in urban environments often have gaps and inaccuracies due to various factors such as sensor malfunctions, environmental conditions, or data transmission errors [164]. Effective MSIF algorithms must compensate for these imperfections by using redundant information from multiple sources to ensure robust decision-making.

- Information Uncertainty: Urban data is frequently affected by noise, inconsistencies, and ambiguities, complicating the task of accurately identifying and differentiating objects or events [165].

Traditional methods may struggle with high uncertainty, highlighting the need for advanced AI techniques to better handle such conditions.
– Conflicting Data: Data collected from various sources may conflict due to duplication, semantic differences, and varying quality. This issue is critical when integrating heterogeneous data sets, necessitating sophisticated conflict resolution techniques [166] such as entity disambiguation and alignment to maintain data integrity.
– Missing Information: With the proliferation of IoT devices in smart sustainable cities, managing large volumes of monitoring data becomes challenging. Missing information can lead to significant inefficiencies in data processing and reduced accuracy in fusion algorithms [167]. Efficient strategies for handling missing data are essential to maintain the reliability of MSIF systems [168].
– Data Heterogeneity: The diversity of data types, including structured, semi-structured, and unstructured data, poses a significant challenge. For instance, textual data requires different processing techniques [162] compared to image or sensor data. Developing methods to integrate these varied data types effectively is crucial for comprehensive information fusion [169].
– Data Correlation: External factors such as environmental noise can cause correlated biases in sensor measurements [170]. If not addressed, these correlations can undermine the confidence in fusion outcomes [171]. Algorithms must be designed to detect and compensate for these correlated interferences to ensure accurate data integration.

b. Data Alignment: To correct calibration discrepancies and address timing issues, aligning sensor data to a unified reference framework is necessary. This procedure, known as sensor data registration [172], is vital for achieving data consistency in fusion processes, especially crucial in real-time contexts where data from multiple sensors might not be synchronized [173].

c. Dynamic Data Fusion: Urban environments are constantly changing, reflecting shifts in traffic density, weather patterns, and power usage [174]. MSIF systems are required to flexibly merge this continuously varying data to provide up-to-date and accurate analyses [175].

d. Scalability: As urban regions expand and more IoT devices are incorporated, there is a significant increase in data volume. MSIF systems must enhance their scalability to cope with this surge efficiently, maintaining performance through robust infrastructure and sophisticated data processing methods that are designed for larger operational scales [176].

e. Security and Privacy: Protecting sensitive data within MSIF systems is imperative. This is achieved through implementing stringent encryption, secure transmission protocols, and methods that preserve data privacy while balancing the need for data to remain accessible and functional [177].

f. Energy Efficiency: With the extensive installation of sensors and fusion nodes, there is an increased focus on the energy consumption in smart cities. It is crucial to create algorithms and systems that are energy-efficient to ensure that MSIF operations can be sustained in smart city contexts [178].

5.2. Future challenges in multisource information fusion

As the implementation of MSIF systems in smart sustainable cities continues to evolve, several future challenges must be anticipated and addressed to ensure the effectiveness and scalability of these systems. These challenges span across technical, application, and implementation domains, each requiring innovative solutions and adaptive strategies.

5.2.1. Application challenges in MSIF

The early stages of MSIF implementation have revealed various challenges, from theoretical constraints and hardware limitations to designing sensor arrays, selecting appropriate algorithms, evaluating systems, and creating interfaces for human–computer interaction [179]. As the prevalence of massive data sets and the Internet of Things (IoT) grows, new challenges emerge across standard MSIF, real-time MSIF, and event-driven MSIF, prompting the need for innovative solutions to manage these complex issues [160]. Hence, these evolving challenges include:

a. General MSIF: As data volumes continue to expand, two key challenges are becoming increasingly critical [180]:

– Data Association: Involves reordering observations to represent the state of monitored targets accurately, crucial for multi-target tracking [181]. This process becomes complex when dealing with multiple targets, necessitating advanced techniques to manage measurement-to-track and track-to-track associations [182].
– Dimensionality Reduction: Essential for managing the vast amounts of data generated in smart sustainable cities [183]. Techniques for reducing data dimensions help save communication bandwidth and computational resources, whether performed at sensor nodes or fusion centers [184].

b. Real-Time MSIF: Real-time data processing is crucial for applications such as traffic management, emergency response, and autonomous vehicles [173]. Future challenges include:

– Dynamic Data Processing: Algorithms must adapt to rapidly changing data streams, such as those in autonomous driving or real-time surveillance systems [185].
– Reduced Time Costs: The constant influx of data requires methods to minimize processing time and handle redundant information efficiently [186].
– Reliability of Data Transmission: Ensuring timely and reliable data transmission despite potential delays and losses is critical for maintaining the performance of real-time MSIF systems [187].

c. Event-Driven MSIF: Primarily used in applications where specific events trigger data collection and processing [188]. Key future challenges include:

– Event-Triggering Strategies [189]: Developing effective strategies to accurately trigger data collection and processing based on relevant events in the urban environment.
– Event Response Speed: Ensuring rapid response to events is crucial for applications such as disaster management and security monitoring [190].
– Temporal Characteristics [191]: Understanding the timing of events and their impact on data fusion processes is essential for real-time application systems, enhancing their ability to manage transient behaviors effectively.

5.2.2. Implementation challenges in MSIF

As MSIF systems scale, the following implementation challenges are expected to intensify:

a. Integration Complexity: Combining data from multiple sources involves complex integration challenges [180]. This includes ensuring compatibility between different data formats, resolving semantic differences, and managing the interactions between various data sources. Developing standardized protocols and middleware solutions can help address these integration complexities.

b. Maintenance and Upgrades: MSIF systems in smart sustainable cities require regular maintenance and updates to handle evolving data sources and changing urban environments [192]. This involves updating algorithms, incorporating new sensor technologies, and ensuring that the system remains robust against new types of data and potential security threats.

c. Interoperability: Ensuring that MSIF systems can interact seamlessly with other urban systems (e.g., transportation, energy, public safety) is crucial for integrated smart city solutions [1]. Interoperability challenges include developing common data standards, ensuring compatibility between different software platforms, and facilitating data exchange across various city departments and stakeholders.

d. User Adoption and Training: Effective implementation of MSIF systems requires user acceptance and proper training. This involves educating city officials and stakeholders on the benefits and functionalities of MSIF, providing user-friendly interfaces [193], and ensuring that users are equipped to interpret and act on the insights provided by the system. These detailed technical challenges highlight the complexity of implementing MSIF in smart sustainable cities, emphasizing the need for continued innovation in algorithm development, data management, and system integration.

6. Practical applications of MSIF: Case studies

The practical application of multisource information fusion (MSIF) techniques in real-world scenarios is critical to understanding its impact on smart sustainable cities. Below are three case studies that illustrate how MSIF is effectively used to address challenges related to data integration, scalability, and the heterogeneity of data sources in various urban contexts.

6.1. Case study 1: Environmental monitoring in sichuan province, China

Environmental monitoring in Sichuan Province, China, is particularly challenging due to the region's complex terrain and climate variability [194]. Traditional methods relying solely on either ground station data or satellite imagery have been insufficient for precise rainfall measurement and environmental assessment. To improve accuracy, a study integrated rainfall data from 161 meteorological stations with satellite estimates from the TRMM 3B42 dataset using MSIF techniques such as linear regression, weighted regression, and Kalman filter fusion. This approach significantly enhanced the accuracy of rainfall measurements, supporting more reliable hydrological modeling, watershed management, and disaster response planning. The success of this integration underscores the importance of MSIF in managing environmental data in regions with complex geographical features.

6.2. Case study 2: Smart grid integration in urban environments

In smart sustainable cities, the integration of diverse data streams, such as those from smart meters, weather stations, and renewable energy sources, is essential for efficient energy management. A case study on a smart distribution power system demonstrated the application of MSIF to ensure data interoperability and real-time processing [109]. By utilizing standardized data formats and advanced distributed processing technologies, the system was able to integrate and process large-scale streaming data effectively, leading to improved grid stability and energy management. This example highlights the critical role of MSIF in enhancing the efficiency and reliability of smart grids, contributing to the sustainability of urban energy systems.

6.3. Case study 3: Urban carrying capacity in Shanghai

Urban carrying capacity is a key consideration in sustainable city planning, particularly in rapidly growing metropolises like Shanghai [195]. A case study in this city utilized MSIF to integrate remote sensing data, GIS data, and urban planning data to assess and enhance urban carrying capacity. The study addressed the challenges of managing heterogeneous and voluminous data by employing a coupling model that analyzed spatial patterns to optimize resource allocation and urban planning. The application of MSIF in this context led to more informed decision-making, ensuring that the city's resources were allocated efficiently to support sustainable urban growth.

7. Future prospects and directions for MSIF in smart sustainable cities

The future of Multisource Information Fusion (MSIF) in smart sustainable cities holds promising advancements through AI integration, autonomous fusion, and interdisciplinary collaboration. These advancements will enhance urban management and decision-making processes, addressing modern urban challenges effectively.

7.1. Advanced AI integration

Machine Learning and AI: Leveraging machine learning algorithms such as Support Vector Machines, K-nearest neighbors, Random Forests, and deep learning models [196,197] can significantly improve MSIF systems [198]. These algorithms handle the complexities and nonlinearities in urban data, enabling more precise and robust data fusion. For example, deep learning can extract intricate features from high-dimensional data, enhancing traffic prediction models in smart cities [199]. Additionally, techniques like the Multi-Source Information Fusion Graph Convolution Network (MIFGCN) [200] are emerging as powerful tools for handling complex spatial–temporal dependencies in data. MIFGCN, for instance, integrates various auxiliary information such as weather and traffic speed to create a dynamic graph that captures evolving relationships between nodes over time, significantly improving prediction accuracy in applications like traffic flow prediction.

AI-Driven Data Fusion: Utilizing AI to dynamically adjust and optimize data fusion processes allows for real-time analysis and decision-making [201]. AI effectively integrates heterogeneous data sources, making MSIF systems more adaptive and intelligent. Reinforcement learning can help systems continuously refine their fusion strategies based on feedback and changing conditions [202]. Another significant advancement is the Multi-Scale Information Fusion-Based Multiple Correlations for Unsupervised Attribute Selection technique [203]. This method addresses the challenge of extracting relevant attributes from large datasets without prior labeling by integrating multi-scale information with fuzzy relations, which enhances the identification of valuable attributes through a comprehensive unsupervised attribute selection algorithm.

Privacy and Security: Ensuring data privacy and security is crucial [204]. Techniques such as federated learning and secure multiparty computation enable effective data fusion while preserving privacy and protecting against cyber-attacks [205]. Federated learning allows multiple entities to collaboratively train models without sharing raw data, maintaining data confidentiality.

7.2. Quantum computing and AI roadmap for MSIF

a. Quantum Computing: Quantum computing represents a paradigm shift in computational power, leveraging qubits that can exist in multiple states simultaneously, thus allowing for parallel processing on a massive scale. In smart sustainable cities, quantum algorithms have the potential to optimize complex urban systems like traffic management, energy distribution, and environmental monitoring by processing vast amounts of data in real-time.

- Application in MSIF: Quantum-enhanced machine learning algorithms can predict and manage urban traffic flows more efficiently than classical methods by analyzing numerous variables simultaneously. Additionally, quantum computing can accelerate real-time environmental monitoring by rapidly processing data from multiple sensors distributed throughout the city, facilitating quicker responses to environmental hazards.

b. AI Advancements: AI technologies, particularly in deep learning and neural networks, are crucial for enhancing MSIF capabilities in smart sustainable cities. These technologies can be integrated into different levels of data fusion, such as feature-level and decision-level fusion, to improve accuracy and reliability.

- Application in MSIF: AI-driven predictive analytics can anticipate environmental hazards by identifying patterns in data from sensors and satellite imagery. Furthermore, AI-powered autonomous systems can optimize resource allocation in real-time, enhancing the efficiency and responsiveness of urban services.

c. Detailed Roadmap: To ensure a successful integration of quantum computing and AI into MSIF for smart sustainable cities, the following steps are proposed:

- Integration Phase: This phase involves the incorporation of AI models into existing MSIF frameworks, with a specific focus on urban applications like traffic management and environmental monitoring. AI algorithms will be developed and integrated to process multisource data and generate actionable insights. This initial step ensures that the foundational AI components are well-aligned with the current infrastructure of smart sustainable cities.

- Scaling with Quantum Computing: As quantum computing technology matures, the next phase involves integrating quantum algorithms into these AI-driven MSIF frameworks. The purpose is to handle more complex and large-scale data fusion tasks, thereby enhancing processing capabilities. This step enables real-time decision-making in dynamic urban environments, capitalizing on quantum computing's ability to analyze vast datasets simultaneously.

- Deployment and Optimization: In this critical phase, the enhanced MSIF systems, now powered by AI and quantum computing, are deployed across various urban sectors. The focus here is on continuous optimization based on real-world feedback and the evolving needs of the city. The goal is to create scalable and adaptive systems that can evolve as the city's demands change, ensuring that the technology remains relevant and effective.

- Long-Term Vision: Finally, the roadmap explores the potential of combining quantum computing and AI into a hybrid model. This phase focuses on further enhancing the efficiency and effectiveness of MSIF systems, laying the groundwork for the next generation of smart sustainable city technologies. This visionary step ensures that the systems are future-proofed and capable of integrating emerging technologies as they develop.

7.3. Autonomous intelligent fusion

Autonomous intelligent fusion involves developing systems that can independently integrate and analyze multisource information. This includes:

Intelligent Data Processing: Creating data with self-learning, self-purification, and proactive capabilities enhances the efficiency of MSIF systems [206,207]. These intelligent data cells preprocess themselves, reducing the load on central processing units. Edge computing processes data at the source, reducing latency and bandwidth usage [208].

Adaptive Models: MSIF models must autonomously adjust and update based on changing data sources and conditions [209]. This adaptability enhances the robustness and accuracy of fusion models in dynamic urban environments. Transfer learning helps models adapt to new data domains without extensive retraining. In this context, the development of A Multi-Source Information Fusion Model for Outlier Detection [210] is particularly relevant, as it introduces an information set (ISet) that captures the uncertainty in information gains. This model uses a total uncertainty measure (TUM) and a fuzzy knowledge measure (FKM) to map fused data to a fuzzy approximation space, improving the precision of outlier detection and enhancing decision-making, prediction, and categorization.

Fusion Methodologies: Developing intelligent fusion methodologies that recommend algorithms based on specific task objectives streamlines the fusion process [211]. These methods should include functions like information filtering, classification, and reasoning for autonomous decision-making [198]. Ensemble methods improve the robustness of fusion outcomes by combining the strengths of multiple algorithms [212].

Dynamic Evaluation: Adapting evaluation criteria dynamically in response to external changes ensures that MSIF systems remain effective and relevant over time [213]. Real-time monitoring and adaptive thresholds help maintain the accuracy and reliability of the fusion process [214].

7.4. Integrated cognitive systems

Machine-to-Machine Cognitive Integration: Future scenarios will see machines with evolving autonomous learning capabilities integrating their cognitive processes, objectives, and decision outcomes into a unified system [215]. This collective intelligence will enable machines to perform complex tasks more efficiently, such as coordinating smart grid operations or managing city-wide transportation networks [216].

Human–Machine Cognitive Integration: Integrating human cognitive processes with machine intelligence will allow humans to leverage machine capabilities for various tasks [217]. Applications include exoskeletons to aid mobility, machines operating in hazardous environments, and virtual learning environments. This integration will enhance human–machine collaboration, making systems more intuitive and effective [218].

7.5. Interdisciplinary collaboration

Collaboration among experts in cognitive science, AI, machine learning, and human–computer interaction is vital for advancing MSIF [219]. Integrating diverse expertise enables the development of innovative solutions to address the complex challenges of smart sustainable cities [220]. For example, insights from behavioral sciences can improve the modeling of human activities and enhance the predictive capabilities of MSIF systems [221].

7.6. Novel research areas

Multi-Modal Data Fusion: Combining data from various modalities (text, images, speech) provides comprehensive insights, especially in applications like public opinion tracking and medical diagnostics [222]. Multi-modal fusion techniques, such as hybrid deep learning models, integrate diverse data types to offer richer and more accurate information [223].

Context-Aware Data Fusion: Incorporating contextual information (location, time, user preferences) into data fusion processes enhances the relevance and accuracy of results, optimizing services like urban transportation and energy management [224]. Context-aware systems leverage spatial–temporal data to provide more precise and actionable insights [225].

Interactive and Explainable Fusion: Developing interactive data fusion systems [226] allows real-time data exploration and analysis, improving decision-making processes. Explainable fusion methods enhance transparency and trust by making the fusion process understandable [145]. Techniques such as visual analytics help users intuitively understand complex data relationships and fusion results [227].

Ontology-Based Data Fusion: Utilizing ontologies to represent and reason about the semantics of data from multiple sources enables

Table 5
Future trends in MSIF for smart sustainable cities.

Future trends	Description	Example applications
Machine learning and AI [198,233]	Leveraging ML algorithms for analyzing high-dimensional data	Self-driving cars for object detection; Medical imaging for diagnosis
Data protection and security [234]	Safeguarding privacy and securing data during fusion processes	Integrating medical records; Analyzing financial data
Multi-modal data fusion [235]	Combining diverse data types for holistic insights	Public opinion analysis; Multimodal medical diagnostics
Federated data fusion [236]	Enabling data integration across decentralized sources	Disease diagnosis; Supply chain logistics optimization
Context-aware data fusion [237]	Incorporating contextual information for accurate results	Smart city infrastructure; Environmental monitoring
Interactive data fusion [238]	Allowing real-time user interaction with fusion algorithms	Live social media analysis; Financial data exploration
Explainable data fusion [239]	Making fusion results clear and understandable	Visualization tools for data analysis; Contribution assessment
Dynamic data fusion [240]	Managing temporal data changes for improved integration	Weather forecasting; Disease progression tracking
Ontology-based data fusion [241]	Utilizing ontologies for accurate data semantics	Disaster detection; Cross-domain search and analysis
Autonomous data fusion [242]	Enabling self-selection and combination of data sources	Space weather monitoring; Social media trend analysis
Collaborative data fusion [243,244]	Facilitating data combination from multiple parties for comprehensive insights	Law enforcement data integration; Supply chain management
Ethical and legal frameworks [245,246]	Developing responsible data use frameworks	Facial recognition guidelines; Medical data sharing policies

more accurate and meaningful data fusion, crucial for applications like disaster detection and cross-domain search [228]. Ontology-based approaches facilitate the integration of heterogeneous data by providing a common framework for understanding different data types [229].

As discussed earlier, MSIF is set to adopt new methods and technologies to handle complex data sources, leading to more autonomous, dynamic, and interactive processes. Key trends include privacy-preserving data fusion [230], multi-modal data fusion [231], and interactive data fusion [232]. These advancements are crucial for enhancing urban management and decision-making in smart sustainable cities. Additionally, Zhou et al. [180] highlight significant advancements in MSIF, focusing on integrating material data from various sources and the role of AI and big data in enhancing fusion processes. They also provide a systematic categorization of fusion techniques, explore the impact of recent technologies, and offer a roadmap for future research. Table 5 provides detailed insights into these trends and their practical applications.

8. Conclusion

This survey presents a comprehensive exploration into the role of multisource information fusion in smart sustainable cities, highlighting the revolutionary impact of our advanced hierarchical and multi-dimensional classification framework. Specifically custom-designed to address the intricate demands of contemporary urban environments, this approach organizes MSIF processes into distinct dimensions, thereby enhancing urban management through both efficiency and effectiveness.

Our framework significantly improves the precision and adaptability of urban systems by focusing on vital aspects such as data reliability, completeness, and the predictive capabilities of infrastructure. By adopting advanced data fusion goals and methodologies, it expertly addresses problematic data scenarios, optimizes resource management, and generates critical insights for strategic urban planning. This strategic deployment not only boosts operational efficiency but also supports proactive management and sustainability initiatives across various urban sectors. However, applying our proposed methodology also uncovers persistent challenges related to scalability, data privacy, and integration with existing infrastructures. Future research should thus concentrate on addressing these obstacles, refining fusion techniques, and enhancing scalability to fully exploit the capabilities of MSIF. Strategic partnerships among technologists, city planners, and government officials are crucial to propel our initiatives. These collaborations are essential for tailoring MSIF technologies to meet the specific and changing conditions of urban landscapes.

In essence, our novel hierarchical and multi-dimensional classification approach marks a pivotal development in the realm of information fusion, establishing a robust base for future innovations in smart city development. By continuously enhancing and adapting these technologies, we aim to stay ahead of the evolving demands of urban dwellers. This commitment to innovation will help us create urban environments that are not only technologically proficient but also socially inclusive and ecologically sustainable, paving the way for more robust and resilient urban ecosystems.

CRediT authorship contribution statement

Houda Orchi: Writing – original draft, Methodology, Investigation, Conceptualization. **Abdoulaye Baniré Diallo:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Halima Elbiaze:** Writing – review & editing, Supervision. **Essaid Sabir:** Writing – review & editing, Supervision. **Mohamed Sadik:** Validation, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: ORCHI Houda reports financial support was provided by University of Quebec in Montreal. ORCHI Houda reports a relationship with University of Quebec in Montreal that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] S.E. Bibri, Data-driven smart sustainable cities of the future: An evidence synthesis approach to a comprehensive state-of-the-art literature review, Sustain. Futures 3 (2021) 100047.

[2] I. UNEP, The weight of cities–resource requirements of future urbanization, 2018.

[3] H.-S. Choi, S.-K. Song, Direction for a transition toward smart sustainable cities based on the diagnosis of smart city plans, Smart Cities 6 (1) (2022) 156–178.

[4] K. Shahid, M. Hassan, A. Husnain, S. Mukhtar, Information and communication technologies for new generation of sustainable smart cities, in: Sustainable Smart Cities-a Vision for Tomorrow, IntechOpen, 2022.

[5] A. Garg, S.P. Maurya, M.B. Laxmi, K.J. Raut, S.S. Yerasuri, R.M. Shinde, Smart cities and smart supply chain: Integration for sustainable urban development, Power Syst. Technol. 48 (1) (2024) 239–254.

[6] T. Alam, R. Gupta, S. Qamar, A. Ullah, Recent applications of artificial intelligence for sustainable development in smart cities, in: Recent Innovations in Artificial Intelligence and Smart Applications, Springer, 2022, pp. 135–154.

[7] Y. Long, X. Zhang, F. Gao, Urban fine management of multisource spatial data fusion based on smart city construction, Math. Probl. Eng. 2021 (2021) 1–10.

[8] H. Yuan, Z. Yin, C. Zhao, Z. Yang, S. Gao, S. Zhao, L. Xu, T. Tan, Y. Fang, Evaluation of smart environmental protection systems and novel UV-oriented solution for integration, resilience, inclusiveness and sustainability, in: 2020 5th International Conference on Universal Village, UV, IEEE, 2020, pp. 1–77.

[9] S. Boopathi, Advancements in machine learning and AI for intelligent systems in drone applications for smart city developments, in: Futuristic E-Governance Security with Deep Learning Applications, IGI Global, 2024, pp. 15–45.

[10] B. Wei, et al., Application of neural network based on multisource information fusion in production cost prediction, Wirel. Commun. Mob. Comput. 2022 (2022).

[11] Z. Qin, H. Liao, L. Chen, L. Zhang, Enterprise performance management following big data analysis technology under multisource information fusion, Secur. Commun. Netw. 2021 (2021) 1–11.

[12] G. Doctor, P. Patel, Air quality monitoring using internet of things (IoT) in smart cities, in: Handbook of Research on Implementation and Deployment of IoT Projects in Smart Cities, IGI Global, 2019, pp. 317–333.

[13] F.A. Almalki, S.H. Alsamhi, R. Sahal, J. Hassan, A. Hawbani, N. Rajput, A. Saif, J. Morgan, J. Breslin, Green IoT for eco-friendly and sustainable smart cities: future directions and opportunities, Mob. Netw. Appl. 28 (1) (2023) 178–202.

[14] K. Pardini, J.J. Rodrigues, O. Diallo, A.K. Das, V.H.C. de Albuquerque, S.A. Kozlov, A smart waste management solution geared towards citizens, Sensors 20 (8) (2020) 2380.

[15] S.M. Shariati, M.H. Ahmadzadegan, et al., Challenges and security issues in cloud computing from two perspectives: Data security and privacy protection, in: 2015 2nd International Conference on Knowledge-Based Engineering and Innovation, KBEI, IEEE, 2015, pp. 1078–1082.

[16] S. Badsha, I. Vakilinia, S. Sengupta, Privacy preserving cyber threat information sharing and learning for cyber defense, in: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference, CCWC, IEEE, 2019, pp. 0708–0714.

[17] H. Su, Z. Wen, X. Sun, X. Yan, Multisource information fusion-based approach diagnosing structural behavior of dam engineering, Struct. Control Health Monit. 25 (2) (2018) e2073.

[18] N. Priyashani, N. Kankanamge, T. Yigitcanlar, Multisource open geospatial big data fusion: Application of the method to demarcate urban agglomeration footprints, Land 12 (2) (2023) 407.

[19] B. Hasan, A.A. Rezaei, J. Marx Gómez, Enhancing privacy in smart cities by facilitating communications between public and private sectors, in: Smart Cities/Smart Regions–Technische, Wirtschaftliche Und Gesellschaftliche Innovationen: Konferenzband Zu Den 10. BUIS-Tagen, Springer, 2019, pp. 123–132.

[20] M. Rahouti, K. Xiong, Y. Xin, Secure software-defined networking communication systems for smart cities: Current status, challenges, and trends, IEEE Access 9 (2021) 12083–12113, http://dx.doi.org/10.1109/ACCESS.2020.3047996.

[21] A. Masood, K. Ahmad, Assessment of municipal solid waste management in jammu city: problems, prospects and solutions, in: Smart Cities—Opportunities and Challenges: Select Proceedings of ICSC 2019, Springer, 2020, pp. 257–275.

[22] A. Harakan, Indonesia's city diplomacy in promoting global sustainability context through urban networks, Ganaya: J. Ilmu Sos. Hum. 6 (3) (2023) 529–538.

[23] O. Kolotouchkina, C.L. Barroso, J.L.M. Sánchez, Smart cities, the digital divide, and people with disabilities, Cities 123 (2022) 103613.

[24] S.-Y. Shin, D. Kim, S.A. Chun, Digital divide in advanced smart city innovations, Sustainability 13 (7) (2021) 4076.

[25] P. Zhang, T. Li, G. Wang, C. Luo, H. Chen, J. Zhang, D. Wang, Z. Yu, Multi-source information fusion based on rough set theory: A review, Inf. Fusion 68 (2021) 85–117.

[26] W. Koch, On digital ethics for artificial intelligence and information fusion in the defense domain, IEEE Aerosp. Electron. Syst. Mag. 36 (7) (2021) 94–111.

[27] H.-X. Li, X.-F. Guo, Research on multi-source information fusion technology, in: International Academic Conference on Frontiers in Social Sciences and Management Innovation, IAFSM 2019, Atlantis Press, 2020, pp. 24–28.

[28] S. Zhenzhen, J. Hongbing, T. Cong, Y. Zhang, A robust Poisson multi-Bernoulli filter for multi-target tracking based on arithmetic average fusion, Chin. J. Aeronaut. 36 (2) (2023) 179–190.

[29] Y. Tianle, M. Run, W. Weili, L. Zhirong, D. Jun, G. Yajuan, Y. Xuefei, Synthetic damage effect assessment through evidential reasoning approach and neural fuzzy inference: Application in ship target, Chin. J. Aeronaut. 35 (8) (2022) 143–157.

[30] J. Jin, W. Zhou, R. Yang, L. Ye, L. Yu, Edge detection guide network for semantic segmentation of remote-sensing images, IEEE Geosci. Remote Sens. Lett. 20 (2023) 1–5.

[31] C. Zhu, F. Xiao, Z. Cao, A generalized rényi divergence for multi-source information fusion with its application in EEG data analysis, Inform. Sci. 605 (2022) 225–243.

[32] P. Song, C. Chen, L. Zhang, Evaluation model of click rate of electronic commerce advertising based on fuzzy genetic algorithm, Mob. Netw. Appl. 27 (3) (2022) 936–945.

[33] M. Wang, Check for updates research on quadrotor UAV path planning optimization based on multi-source information fusion technology of ant colony optimization algorithm, in: Communications, Signal Processing, and Systems: Proceedings of the 11th International Conference on Communications, Signal Processing, and Systems, Vol. 1, Springer Nature Amsterdam, 2023, p. 162.

[34] W. Xiwei, X. Bing, W. Cihang, G. Yiming, L. Lingwei, Factor graph based navigation and positioning for control system design: A review, Chin. J. Aeronaut. 35 (5) (2022) 25–39.

[35] J. Deng, H. Liu, H. Fang, S. Shao, D. Wang, Y. Hou, D. Chen, M. Tang, MgNet: A fault diagnosis approach for multi-bearing system based on auxiliary bearing and multi-granularity information fusion, Mech. Syst. Signal Process. 193 (2023) 110253.

[36] B.P.L. Lau, S.H. Marakkalage, Y. Zhou, N.U. Hassan, C. Yuen, M. Zhang, U.-X. Tan, A survey of data fusion in smart city applications, Inf. Fusion 52 (2019) 357–374.

[37] M.A. Fadhel, A.M. Duhaim, A. Saihood, A. Sewify, M.N. Al-Hamadani, A. Albahri, L. Alzubaidi, A. Gupta, S. Mirjalili, Y. Gu, Comprehensive systematic review of information fusion methods in smart cities and urban environments, Inf. Fusion (2024) 102317.

[38] X. Li, Q. Yu, B. Alzahrani, A. Barnawi, A. Alhindi, D. Alghazzawi, Y. Miao, Data fusion for intelligent crowd monitoring and management systems: A survey, IEEE Access 9 (2021) 47069–47083.

[39] C. Ounoughi, S.B. Yahia, Data fusion for ITS: A systematic literature review, Inf. Fusion 89 (2023) 267–291.

[40] N.-E. El Faouzi, H. Leung, A. Kurian, Data fusion in intelligent transportation systems: Progress and challenges–A survey, Inf. Fusion 12 (1) (2011) 4–10.

[41] B. Esmaeilian, B. Wang, K. Lewis, F. Duarte, C. Ratti, S. Behdad, The future of waste management in smart and sustainable cities: A review and concept paper, Waste Manage. 81 (2018) 177–195.

[42] X. Qin, Y. Gu, Data fusion in the internet of things, Procedia Eng. 15 (2011) 3023–3026.

[43] S. Hassani, U. Dackermann, M. Mousavi, J. Li, A systematic review of data fusion techniques for optimized structural health monitoring, Inf. Fusion (2023) 102136.

[44] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, B. Hu, Feature-level fusion approaches based on multimodal EEG data for depression recognition, Inf. Fusion 59 (2020) 127–138.

[45] Y. Wei, D. Wu, J. Terpenny, Decision-level data fusion in quality control and predictive maintenance, IEEE Trans. Autom. Sci. Eng. 18 (1) (2020) 184–194.

[46] J.-W. Benjamins, M.W. Yeung, A.E. Reyes-Quintero, B. Ruijsink, P. van der Harst, L.E. Juarez-Orozco, Hybrid cardiac imaging: The role of machine learning and artificial intelligence, in: Hybrid Cardiac Imaging for Clinical Decision-Making: From Diagnosis to Prognosis, Springer, 2022, pp. 203–222.

[47] Y. Zhang, A.V. Malawade, X. Zhang, Y. Li, D. Seong, M.A. Al Faruque, S. Huang, CARMA: Context-aware runtime reconfiguration for energy-efficient sensor fusion, in: 2023 IEEE/ACM International Symposium on Low Power Electronics and Design, ISLPED, IEEE, 2023, pp. 1–6.

[48] E. Blasch, T. Pham, C.-Y. Chong, W. Koch, H. Leung, D. Braines, T. Abdelzaher, Machine learning/artificial intelligence for sensor data fusion–opportunities and challenges, IEEE Aerosp. Electron. Syst. Mag. 36 (7) (2021) 80–93.

[49] H. Orchi, M. Sadik, M. Khaldoun, A general survey on plants disease detection using image processing, deep transfer learning and machine learning techniques, in: Ubiquitous Networking: 7th International Symposium, UNet 2021, Virtual Event, May 19–22, 2021, Revised Selected Papers 7, Springer, 2021, pp. 210–224.

[50] L. Comment, B.A. Coull, C. Zigler, L. Valeri, Bayesian data fusion: Probabilistic sensitivity analysis for unmeasured confounding using informative priors based on secondary data, Biometrics 78 (2) (2022) 730–741.

[51] E. Blasch, I. Kadar, J. Salerno, M.M. Kokar, S. Das, G.M. Powell, D.D. Corkill, E.H. Ruspini, Issues and challenges of knowledge representation and reasoning methods in situation assessment (level 2 fusion), Signal Process. Sens. Fusion Target Recognit. XV 6235 (2006) 355–368.

[52] J. Llinas, C. Bowman, G. Rogova, A. Steinberg, E. Waltz, F. White, et al., Revisiting the JDL data fusion model II, in: Proceedings of the Seventh International Conference on Information Fusion, Vol. 2, FOI Stockholm, Sweden, 2004, pp. 1218–1230.
[53] B.V. Dasarathy, Decision Fusion, vol. 1994, IEEE Computer Society Press Los Alamitos, 1994.
[54] B.V. Dasarathy, Decision fusion strategies in multisensor environments, IEEE Trans. Syst. Man Cybern. 21 (5) (1991) 1140–1154.
[55] I.R. Goodman, R.P. Mahler, H.T. Nguyen, Mathematics of Data Fusion, vol. 37, Springer Science & Business Media, 2013.
[56] C. Bowman, C. Morefield, Multisensor fusion of target attributes and kinematics, in: 1980 19th IEEE Conference on Decision and Control Including the Symposium on Adaptive Processes, IEEE, 1980, pp. 837–839.
[57] R.C. Luo, M.G. Kay, Multisensor integration and fusion: issues and approaches, in: Sensor Fusion, Vol. 931, SPIE, 1988, pp. 42–49.
[58] L.-F. Pau, Sensor data fusion, J. Intell. Robot. Syst. 1 (1988) 103–116.
[59] S.C. Thomopoulos, Sensor integration and data fusion, J. Robot. Syst. 7 (3) (1990) 337–372.
[60] C. Harris, A. Bailey, T. Dodd, Multi-sensor data fusion in defence and aerospace, Aeronaut. J. 102 (1015) (1998) 229–244.
[61] M. Bedworth, J. O'Brien, The omnibus model: a new model of data fusion? IEEE Aerosp. Electron. Syst. Mag. 15 (4) (2000) 30–36.
[62] J. Schoess, G. Castore, A distributed sensor architecture for advanced aerospace systems, in: Sensor Fusion, Vol. 931, SPIE, 1988, pp. 74–87.
[63] S. Thrun, Probabilistic robotics, Commun. ACM 45 (3) (2002) 52–57.
[64] H. Durrant-Whyte, T.C. Henderson, Multisensor data fusion, in: Springer Handbook of Robotics, Springer, 2016, pp. 867–896.
[65] D.L. Hall, S.A. McMullen, Mathematical Techniques in Multisensor Data Fusion, Artech House, 2004.
[66] M.I. Ribeiro, Kalman and extended kalman filters: Concept, derivation and properties, Inst. Syst. Robot. 43 (46) (2004) 3736–3741.
[67] E.A. Wan, R. Van der Merwe, Kalman filtering and neural networks, Unscented Kalman Filter 221 (2001) 277.
[68] L.D. Stone, R.L. Streit, T.L. Corwin, K.L. Bell, Bayesian Multiple Target Tracking, Artech House, 2013.
[69] M.K. Cowles, B.P. Carlin, Markov chain Monte Carlo convergence diagnostics: a comparative review, J. Am. Statist. Assoc. 91 (434) (1996) 883–904.
[70] A.P. Dempster, A generalization of Bayesian inference, J. R. Stat. Soc. Ser. B Stat. Methodol. 30 (2) (1968) 205–232.
[71] A. Campagner, D. Ciucci, T. Denœux, Belief functions and rough sets: Survey and new insights, Internat. J. Approx. Reason. 143 (2022) 192–215.
[72] K. Zhao, L. Li, Z. Chen, R. Sun, G. Yuan, J. Li, A survey: Optimization and applications of evidence fusion algorithm based on Dempster–Shafer theory, Appl. Soft Comput. 124 (2022) 109075.
[73] J. Dezert, Advances and Applications of Dsmt for Information Fusion: Collected Works, American Research Press, 2009.
[74] T.D. Garvey, J.D. Lowrance, M.A. Fischler, An inference technique for integrating knowledge from disparate sources, in: IJCAI, Vol. 7, 1981.
[75] F. Pichon, D. Dubois, T. Denœux, Quality of information sources in information fusion, in: Information Quality in Information Fusion and Decision Making, Springer, 2019, pp. 31–49.
[76] Y. Donga, X. Lia, J. Dezertc, M.O. Khyamd, M. Noor-A-Rahime, S.S. Gef, DSmT-based fusion strategy for human activity recognition in body sensor networks, Adv. Appl. DSmT Inf. Fusion (2023) 419.
[77] Y. Dong, X. Li, J. Dezert, R. Zhou, C. Zhu, L. Cao, M.O. Khyam, S.S. Ge, Multisource weighted domain adaptation with evidential reasoning for activity recognition, IEEE Trans. Ind. Inform. 19 (4) (2022) 5530–5542.
[78] F. Xiao, GEJS: A generalized evidential divergence measure for multisource information fusion, IEEE Trans. Syst. Man Cybern. Syst. 53 (4) (2022) 2246–2258.
[79] Y. Deng, D numbers theory: a generalization of Dempster–Shafer evidence theory, 2014, arXiv preprint arXiv:1405.3175.
[80] W. Pedrycz, A. Gacek, X. Wang, Aggregation of order-2 fuzzy sets, IEEE Trans. Fuzzy Syst. 29 (11) (2020) 3570–3575.
[81] E. Zarei, F. Khan, R. Abbassi, Importance of human reliability in process operation: A critical analysis, Reliab. Eng. Syst. Saf. 211 (2021) 107607.
[82] F. Xiao, Z. Cao, C.-T. Lin, A complex weighted discounting multisource information fusion with its application in pattern classification, IEEE Trans. Knowl. Data Eng. (2022).
[83] F. Xiao, W. Pedrycz, Negation of the quantum mass function for multisource quantum information fusion with its application to pattern classification, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2) (2022) 2054–2070.
[84] B.R. Bracio, W. Horn, D.P. Moller, Sensor fusion in biomedical systems, in: Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No. 97CH36136), Vol. 3, IEEE, 1997, pp. 1387–1390.
[85] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (3) (1965) 338–353.
[86] D.J. Dubois, Fuzzy Sets and Systems: Theory and Applications, vol. 144, Academic Press, 1980.
[87] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic, vol. 4, Prentice Hall, New Jersey, 1995.
[88] H.-J. Zimmermann, Fuzzy Set Theory—and its Applications, Springer Science & Business Media, 2011.
[89] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst. Man Cybern. (1) (1973) 28–44.
[90] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, vol. 9, Springer Science & Business Media, 2012.
[91] Y. Yao, Information granulation and rough set approximation, Int. J. Intell. Syst. 16 (1) (2001) 87–104.
[92] R. Jensen, Q. Shen, Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches, IEEE Trans. Knowl. Data Eng. 16 (12) (2004) 1457–1471.
[93] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: A tutorial, in: Rough Fuzzy Hybridization: a New Trend in Decision-Making, Springer, 1999, pp. 3–98.
[94] D. Wang, T. Li, W. Huang, Z. Luo, P. Deng, P. Zhang, M. Ma, A multi-view clustering algorithm based on deep semi-NMF, Inf. Fusion 99 (2023) 101884.
[95] H. He, F. Xiao, A novel quantum Dempster's rule of combination for pattern classification, Inform. Sci. 671 (2024) 120617.
[96] Y. Huang, F. Xiao, Fractal belief Jensen–Shannon divergence-based multi-source information fusion for pattern classification, Eng. Appl. Artif. Intell. 126 (2023) 107048.
[97] Y. Huang, F. Xiao, Z. Cao, C.-T. Lin, Higher order fractal belief Rényi divergence with its applications in pattern classification, IEEE Trans. Pattern Anal. Mach. Intell. (2023).
[98] X.L. Dong, L. Berti-Equille, D. Srivastava, Data fusion: resolving conflicts from multiple sources, in: Handbook of Data Quality: Research and Practice, Springer, 2013, pp. 293–318.
[99] X. Wang, L.-P. Huang, Y. Zhang, X.-H. Xu, J.-Q. Chen, A solution of data inconsistencies in data integration—Designed for pervasive computing environment, J. Comput. Sci. Tech. 25 (3) (2010) 499–508.
[100] R. Gutiérrez, V. Rampérez, H. Paggi, J.A. Lara, J. Soriano, On the use of information fusion techniques to improve information quality: Taxonomy, opportunities and challenges, Inf. Fusion 78 (2022) 102–137.
[101] M.A. Becerra, C. Tobón, A.E. Castro-Ospina, D.H. Peluffo-Ordóñez, Information quality assessment for data fusion systems, Data 6 (6) (2021) 60.
[102] N. Brierley, T. Tippetts, P. Cawley, Data fusion for improving the reliability of automated non-destructive inspection, in: AIP Conference Proceedings, Vol. 1511, American Institute of Physics, 2013, pp. 1781–1788.
[103] S. Baskar, R. Selvaraj, V.M. Kuthadi, P.M. Shakeel, Attribute-based data fusion for designing a rational trust model for improving the service reliability of internet of things assisted applications in smart cities, Soft Comput. 25 (18) (2021) 12275–12289.
[104] F.A. Silva, I. Fé, F. Silva, T.A. Nguyen, Quantifying the impact of resource redundancy on smart city system dependability: a model-driven approach, Cluster Comput. (2024) 1–21.
[105] X. Zou, Y. Yan, X. Hao, Y. Hu, H. Wen, E. Liu, J. Zhang, Y. Li, T. Li, Y. Zheng, et al., Deep learning for cross-domain data fusion in urban computing: Taxonomy, advances, and outlook, 2024, arXiv preprint arXiv:2402.19348.
[106] S. Krishnan Ravi, Y. Comlek, W. Chen, A. Pathak, V. Gupta, R. Umretiya, A. Hoffman, G. Pilania, P. Pandita, S. Ghosh, et al., Interpretable multi-source data fusion through latent variable Gaussian process, 2024, arXiv e-prints, arXiv–2402.
[107] Q. Yuan, Y. Pi, L. Kou, F. Zhang, Y. Li, Z. Zhang, Multi-source data processing and fusion method for power distribution internet of things based on edge intelligence, Front. Energy Res. 10 (2022) 891867.
[108] R.S. Ferraz, R.S. Ferraz, A.C. Rueda-Medina, Multi-objective optimization approach for allocation and sizing of distributed energy resources preserving the protection scheme in distribution networks, J. Control Autom. Electr. Syst. 34 (5) (2023) 1080–1092.
[109] X. He, H. Dong, W. Yang, W. Li, Multi-source information fusion technology and its application in smart distribution power system, Sustainability 15 (7) (2023) 6170.
[110] A. Samarakkody, D. Amaratunga, R. Haigh, Technological innovations for enhancing disaster resilience in smart cities: a comprehensive urban Scholar's analysis, Sustainability 15 (15) (2023) 12036.
[111] R.-C. Chen, C. Dewi, S.-W. Huang, R.E. Caraka, Selecting critical features for data classification based on machine learning methods, J. Big Data 7 (1) (2020) 52.
[112] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27.
[113] Y. Bar-Shalom, F. Daum, J. Huang, The probabilistic data association filter, IEEE Control Syst. Mag. 29 (6) (2009) 82–100.
[114] J.P. Shaffer, Multiple hypothesis testing, Annu. Rev. Psychol. 46 (1) (1995) 561–584.
[115] G. Welch, G. Bishop, et al., An introduction to the Kalman filter, 1995.
[116] I. Lourenço, R. Mattila, C.R. Rojas, X. Hu, B. Wahlberg, Hidden Markov models: Inverse filtering, belief estimation and privacy protection, J. Syst. Sci. Complex. 34 (2021) 1801–1820.

[117] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman Filter: Particle Filters for Tracking Applications, Artech House, 2003.
[118] I.J. Myung, Tutorial on maximum likelihood estimation, J. Math. Psychol. 47 (1) (2003) 90–100.
[119] J.K. Uhlmann, Covariance consistency methods for fault-tolerant distributed data fusion, Inf. Fusion 4 (3) (2003) 201–215.
[120] J. Teremranova, A. Mutule, Smart approach to management of energy resources in smart cities: Evaluation of models and methods, in: Smart and Sustainable Planning for Cities and Regions: Results of SSPCR 2019 3, Springer, 2021, pp. 3–16.
[121] P. Mishra, G. Singh, Energy management of sustainable smart cities using internet-of-energy, in: Sustainable Smart Cities: Enabling Technologies, Energy Trends and Potential Applications, Springer, 2023, pp. 143–173.
[122] H. Wu, M. Siegel, R. Stiefelhagen, J. Yang, Sensor fusion using Dempster-Shafer theory [for context-aware hci], in: IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No. 00CH37276), Vol. 1, IEEE, 2002, pp. 7–12.
[123] F. Herrera, E. Herrera-Viedma, L. Martınez, A fusion approach for managing multi-granularity linguistic term sets in decision making, Fuzzy Sets and Systems 114 (1) (2000) 43–58.
[124] G.E. Box, G.C. Tiao, Bayesian Inference in Statistical Analysis, John Wiley & Sons, 2011.
[125] M. Yu, H. Xu, F. Zhou, S. Xu, H. Yin, A deep-learning-based multimodal data fusion framework for urban region function recognition, ISPRS Int. J. Geo-Inf. 12 (12) (2023) 468.
[126] T.T. Cai, R. Ma, Theoretical foundations of t-sne for visualizing high-dimensional clustered data, J. Mach. Learn. Res. 23 (301) (2022) 1–54.
[127] I.T. Jolliffe, Principal Component Analysis for Special Types of Data, Springer, 2002.
[128] W.H. Hsu, A position paper on statistical inference techniques which integrate neural network and Bayesian network models, in: Proceedings of International Conference on Neural Networks, ICNN'97, Vol. 3, IEEE, 1997, pp. 1972–1977.
[129] F. Emmert-Streib, M. Dehmer, Understanding statistical hypothesis testing: The logic of statistical inference, Mach. Learn. Knowl. Extract. 1 (3) (2019) 945–962.
[130] C.-T. Fan, Y.-K. Wang, C.-R. Huang, Heterogeneous information fusion and visualization for a large-scale intelligent video surveillance system, IEEE Trans. Syst. Man Cybern. Syst. 47 (4) (2016) 593–604.
[131] C. Ware, Information Visualization: Perception for Design, Morgan Kaufmann, 2019.
[132] B.V. Dasarathy, Sensor fusion potential exploitation-innovative architectures and illustrative applications, Proc. IEEE 85 (1) (1997) 24–38.
[133] W. Hurst, B. Tekinerdogan, B. Kotze, Perceptions on smart gas meters in smart cities for reducing the carbon footprint, Smart Cities 3 (4) (2020) 1173–1186.
[134] T. Breur, Data analysis across various media: Data fusion, direct marketing, clickstream data and social media, J. Direct Data Digit. Mark. Pract. 13 (2011) 95–105.
[135] J.A. Balazs, J.D. Velásquez, Opinion mining and information fusion: a survey, Inf. Fusion 27 (2016) 95–110.
[136] S.H. Marakkalage, S. Sarica, B.P.L. Lau, S.K. Viswanath, T. Balasubramaniam, C. Yuen, B. Yuen, J. Luo, R. Nayak, Understanding the lifestyle of older population: Mobile crowdsensing approach, IEEE Trans. Comput. Soc. Syst. 6 (1) (2018) 82–95.
[137] Y. Ma, Y. Zeng, X. Ren, N. Zhong, User interests modeling based on multi-source personal information fusion and semantic reasoning, in: Active Media Technology: 7th International Conference, AMT 2011, Lanzhou, China, September 7-9, 2011. Proceedings 7, Springer, 2011, pp. 195–205.
[138] L. You, B. Tuncer, H. Xing, Harnessing multi-source data about public sentiments and activities for informed design, IEEE Trans. Knowl. Data Eng. 31 (2) (2018) 343–356.
[139] Q. Mei, M. Gül, N. Shirzad-Ghaleroudkhani, Towards smart cities: crowdsensing-based monitoring of transportation infrastructure using in-traffic vehicles, J. Civ. Struct. Health Monit. 10 (4) (2020) 653–665.
[140] L. Jayasinghe, N. Wijerathne, C. Yuen, M. Zhang, Feature learning and analysis for cleanliness classification in restrooms, IEEE Access 7 (2019) 14871–14882.
[141] W. Tushar, N. Wijerathne, W.-T. Li, C. Yuen, H.V. Poor, T.K. Saha, K.L. Wood, Internet of things for green building management: disruptive innovations through low-cost sensor technology and artificial intelligence, IEEE Signal Process. Mag. 35 (5) (2018) 100–110.
[142] O. Hakimi, H. Liu, O. Abudayyeh, A. Houshyar, M. Almatared, A. Alhawiti, Data fusion for smart civil infrastructure management: A conceptual digital twin framework, Buildings 13 (11) (2023) 2725.
[143] Y. Li, S. Chen, K. Hwang, X. Ji, Z. Lei, Y. Zhu, F. Ye, M. Liu, Spatio-temporal data fusion techniques for modeling digital twin city, Geo-Spat. Inf. Sci. (2024) 1–24.
[144] S. Izumi, S.-i. Azuma, Real-time pricing by data fusion on networks, IEEE Trans. Ind. Inform. 14 (3) (2017) 1175–1185.
[145] S. Sinha, P. Vashisht, Explainable data fusion on edge: Challenges and opportunities, in: Explainable Edge AI: A Futuristic Computing Perspective, Springer, 2022, pp. 117–138.
[146] G. Mujica, R. Rodriguez-Zurrunero, M.R. Wilby, J. Portilla, A.B. Rodríguez González, A. Araujo, T. Riesgo, J.J. Vinagre Diaz, Edge and fog computing platform for data fusion of complex heterogeneous sensors, Sensors 18 (11) (2018) 3630.
[147] F.J. Valente, J.P. Morijo, K.C.T. Vivaldini, L.C. Trevelin, Fog-based data fusion for heterogeneous iot sensor networks: a real implementation, in: 2019 15th International Conference on Network and Service Management, CNSM, IEEE, 2019, pp. 1–5.
[148] B. Tai, X. Li, L. Yang, Y. Miao, W. Lin, C. Yan, Cloud computing-aided multi-type data fusion with correlation for education, Wirel. Netw. (2022) 1–12.
[149] L. Fan, L. Zhang, Multi-system fusion based on deep neural network and cloud edge computing and its application in intelligent manufacturing, Neural Comput. Appl. 34 (5) (2022) 3411–3420.
[150] C. Wang, M. Zhang, F. Shi, P. Xue, Y. Li, A hybrid multimodal data fusion-based method for identifying gambling websites, Electronics 11 (16) (2022) 2489.
[151] Y. Liu, Multi-source heterogeneous data fusion based on perceptual semantics in narrow-band internet of things, Pers. Ubiquitous Comput. 23 (3) (2019) 413–420.
[152] G. Sahu, O. Vechtomova, Adaptive fusion techniques for multimodal data, 2019, arXiv preprint arXiv:1911.03821.
[153] H. Orchi, M. Sadik, M. Khaldoun, A novel hybrid deep learning model for crop disease detection using began, in: International Symposium on Ubiquitous Networking, Springer, 2022, pp. 267–283.
[154] C. Qin, B. Guo, Y. Zhang, Y. Shen, The dynamic fusion representation of multi-source fuzzy data, Appl. Intell. 53 (22) (2023) 27226–27248.
[155] R. Ahmad, I. Alsmadi, Data fusion and network intrusion detection systems, Cluster Comput. (2024) 1–27.
[156] K. Zhang, Y. Sun, R. Wang, H. Li, X. Hu, Multiple fusion adaptation: A strong framework for unsupervised semantic segmentation adaptation, 2021, arXiv preprint arXiv:2112.00295.
[157] F. Zhu, Z. Xu, X. Zhang, Y. Zhang, W. Chen, X. Zhang, On state estimation in multi-sensor fusion navigation: Optimization and filtering, 2024, arXiv preprint arXiv:2401.05836.
[158] A. De Paola, P. Ferraro, S. Gaglio, G. Lo Re, Context-awareness for multi-sensor data fusion in smart environments, in: AI* IA 2016 Advances in Artificial Intelligence: XVth International Conference of the Italian Association for Artificial Intelligence, Genova, Italy, November 29–December 1, 2016, Proceedings XV, Springer, 2016, pp. 377–391.
[159] B. Kaci, L. Benjamin, S. Walter, A fault tolerant architecture for data fusion targeting hardware and software faults, in: 2014 IEEE 20th Pacific Rim International Symposium on Dependable Computing, IEEE, 2014, pp. 1–10.
[160] K. Cai, W. Xu, An efficient multi-source information fusion approach for dynamic interval-valued data via fuzzy approximate conditional entropy, Int. J. Mach. Learn. Cybern. (2024) 1–27.
[161] X. Zhang, Q. Wu, X. Wang, Y. Chen, Research on multi-source data fusion technology under power cloud platform, in: Advances in Intelligent Systems and Interactive Applications: Proceedings of the 4th International Conference on Intelligent, Interactive Systems and Applications (IISA2019) 4, Springer, 2020, pp. 554–559.
[162] I.A. Al-Baltah, A.A.A. Ghani, G.M. Al-Gomaei, F.H. Abdulrazzak, A.A.A. Kharusi, A scalable semantic data fusion framework for heterogeneous sensors data, J. Ambient Intell. Humaniz. Comput. 14 (5) (2023) 5047–5066.
[163] P. Tanaka, S. Sapra, N. Laptev, Scalable data classification for security and privacy, 2020, arXiv preprint arXiv:2006.14109.
[164] B. Agbo, H. Al-Aqrabi, R. Hill, T. Alsboui, Missing data imputation in the internet of things sensor networks, Future Internet 14 (5) (2022) 143.
[165] K.K. Kumar, E. Ramaraj, P. Geetha, Multi-sensor data fusion for an efficient object tracking in internet of things (IoT), Appl. Nanosci. 13 (2) (2023) 1355–1365.
[166] A. Kulmukhametov, A. Rauber, C. Becker, Improving data quality in large-scale repositories through conflict resolution, Int. J. Digit. Libr. 22 (4) (2021) 365–383.
[167] Y. Chen, S. Li, J. Yao, Processing missing information in big data environment, in: Data Mining and Big Data: Third International Conference, DMBD 2018, Shanghai, China, June 17–22, 2018, Proceedings 3, Springer, 2018, pp. 638–648.
[168] F.A. Adnan, K.R. Jamaludin, W.Z.A. Wan Muhamad, S. Miskon, A review of the current publication trends on missing data imputation over three decades: direction and future research, Neural Comput. Appl. 34 (21) (2022) 18325–18340.
[169] L. Wang, Heterogeneous data and big data analytics, Autom. Control Inf. Sci. 3 (1) (2017) 8–15.
[170] V.G. Biju, A.-M. Schmitt, B. Engelmann, Assessing the influence of sensor-induced noise on machine-learning-based changeover detection in CNC machines, Sensors 24 (2) (2024) 330.
[171] M. Abu Bakr, S. Lee, Distributed multisensor data fusion under unknown correlation and data inconsistency, Sensors 17 (11) (2017) 2472.
[172] A. Alempijevic, S. Kodagoda, G. Dissanayake, Sensor registration for robotic applications, in: Field and Service Robotics: Results of the 6th International Conference, Springer, 2008, pp. 233–242.

[173] N. Senel, K. Kefferpütz, K. Doycheva, G. Elger, Multi-sensor data fusion for real-time multi-object tracking, Processes 11 (2) (2023) 501.

[174] H. Ma, X. Qin, Y. Jia, J. Zhou, Dynamic spatio-temporal graph fusion convolutional network for urban traffic prediction, Appl. Sci. 13 (16) (2023) 9304.

[175] Y. Huang, T. Li, C. Luo, H. Fujita, S.-J. Horng, Dynamic fusion of multisource interval-valued data by fuzzy granulation, IEEE Trans. Fuzzy Syst. 26 (6) (2018) 3403–3417.

[176] N. Nithiyanandam, M. Rajesh, R. Sitharthan, D. Shanmuga Sundar, K. Vengatesan, K. Madurakavi, Optimization of performance and scalability measures across cloud based IoT applications with efficient scheduling approach, Int. J. Wirel. Inf. Netw. 29 (4) (2022) 442–453.

[177] H. Liu, X. Tang, T. Xu, J. He, A survey of homogeneous and heterogeneous multi-source information fusion based on rough set theory, in: International Artificial Intelligence Conference, Springer, 2023, pp. 235–246.

[178] V. Isanbaev, R. Baños, C. Gil, M. Gil, F. Martínez, A. Alcayde, Towards energy efficiency in microgrids for smart sustainable cities, in: Sustainable Smart Cities and Territories International Conference, Springer, 2023, pp. 254–265.

[179] B. Xiang, Y. Li, G. Chen, X. Liu, W. Yang, Multi-source integrated fusion for surface measurement, Int. J. Adv. Manuf. Technol. 109 (2020) 1815–1823.

[180] J. Zhou, X. Hong, P. Jin, Information fusion for multi-source material data: progress and challenges, Appl. Sci. 9 (17) (2019) 3473.

[181] X. Wang, Y. Wang, J. Yang, X. Jia, L. Li, W. Ding, F.-Y. Wang, The survey on multi-source data fusion in cyber-physical-social systems: Foundational infrastructure for industrial metaverses and industries 5.0, Inf. Fusion (2024) 102321.

[182] H.H. Sönmez, A.K. Hocaoğlu, Asynchronous track-to-track association algorithm based on reference topology feature, Signal Image Video Process. (2022) 1–8.

[183] K.R. Varshney, A.S. Willsky, Learning dimensionality-reduced classifiers for information fusion, in: 2009 12th International Conference on Information Fusion, IEEE, 2009, pp. 1881–1888.

[184] A. Bucci, L. He, Z. Liu, Combining dimensionality reduction methods with neural networks for realized volatility forecasting, Ann. Oper. Res. (2023) 1–29.

[185] Y.-b. Meng, T.-y. Li, G.-h. Liu, S.-j. Xu, T. Ji, Real-time dynamic estimation of occupancy load and an air-conditioning predictive control method based on image information fusion, Build. Environ. 173 (2020) 106741.

[186] Z. Lian, S. Shao, C. Huang, A real time face tracking system based on multiple information fusion, Multimedia Tools Appl. 79 (23) (2020) 16751–16769.

[187] Y. Jin, Z. Qian, G. Sun, A real-time multimedia streaming transmission control mechanism based on edge cloud computing and opportunistic approximation optimization, Multimedia Tools Appl. 78 (7) (2019) 8911–8926.

[188] X. Du, H. Liu, H. Yu, Event-triggered robust fusion estimation for multi-sensor time-delay systems with packet drops, Appl. Sci. 13 (15) (2023) 8778.

[189] Z. Hou, Z. Zhou, H. Yuan, W. Wang, J. Wang, Z. Xu, Adaptive event-triggered consensus of multi-agent systems in sense of asymptotic convergence, Sensors 24 (2) (2024) 339.

[190] R. Damaševičius, N. Bacanin, S. Misra, From sensors to safety: Internet of emergency services (IoES) for emergency response and disaster management, J. Sens. Actuat. Netw. 12 (3) (2023) 41.

[191] J. Tang, J. Zeng, L. Zhang, R. Zhang, J. Li, X. Li, J. Zou, Y. Zeng, Z. Xu, Q. Wang, et al., A modified flexible spatiotemporal data fusion model, Front. Earth Sci. 14 (2020) 601–614.

[192] S.E. Bibri, A. Alexandre, A. Sharifi, J. Krogstie, Environmentally sustainable smart cities and their converging AI, IoT, and big data technologies and solutions: an integrated approach to an extensive literature review, Energy Inform. 6 (1) (2023) 9.

[193] Y. Chen, Y. Jiang, H. Liu, Analysis method of app software user experience based on multisource information fusion, Int. J. Semant. Web Inf. Syst. (IJSWIS) 19 (1) (2023) 1–22.

[194] Z. Duan, Y. Ren, X. Liu, H. Lei, X. Hua, X. Shu, L. Zhou, A comprehensive comparison of data fusion approaches to multi-source precipitation observations: A case study in Sichuan province, China, Environ. Monit. Assess. 194 (6) (2022) 422.

[195] X. Cao, Y. Shi, L. Zhou, Research on urban carrying capacity based on multisource data fusion—a case study of Shanghai, Remote Sens. 13 (14) (2021) 2695.

[196] D. Wang, T. Li, P. Deng, Z. Luo, P. Zhang, K. Liu, W. Huang, DNSRF: Deep network-based semi-NMF representation framework, ACM Trans. Intell. Syst. Technol. (2024).

[197] D. Wang, T. Li, P. Deng, F. Zhang, W. Huang, P. Zhang, J. Liu, A generalized deep learning clustering algorithm based on non-negative matrix factorization, ACM Trans. Knowl. Discov. Data 17 (7) (2023) 1–20.

[198] M. Sharma, P. Kushwaha, P. Kumari, P. Kumari, R. Yadav, Machine learning techniques in data fusion: A review, in: International Conference on Communication and Intelligent Systems, Springer, 2022, pp. 391–405.

[199] M.W. Ei Leen, N.H.A. Jafry, N.M. Salleh, H. Hwang, N.A. Jalil, Mitigating traffic congestion in smart and sustainable cities using machine learning: A review, in: International Conference on Computational Science and Its Applications, Springer, 2023, pp. 321–331.

[200] Q. Li, P. Xu, D. He, Y. Wu, H. Tan, X. Yang, Multi-source information fusion graph convolution network for traffic flow prediction, Expert Syst. Appl. 252 (2024) 124288.

[201] S.A. Sayed, Y. Abdel-Hamid, H.A. Hefny, Artificial intelligence-based traffic flow prediction: a comprehensive review, J. Electr. Syst. Inf. Technol. 10 (1) (2023) 13.

[202] I. Gligorea, M. Cioca, R. Oancea, A.-T. Gorski, H. Gorski, P. Tudorache, Adaptive learning using artificial intelligence in e-learning: A literature review, Educ. Sci. 13 (12) (2023) 1216.

[203] P. Zhang, D. Wang, Z. Yu, Y. Zhang, T. Jiang, T. Li, A multi-scale information fusion-based multiple correlations for unsupervised attribute selection, Inf. Fusion 106 (2024) 102276.

[204] X. Zhang, The privacy protection of IoT data fusion and sensing detection under artificial intelligence technology, Wirel. Pers. Commun. (2024) 1–17.

[205] K. Sameera, K. Rafidha Rehiman, P. Vinod, A privacy preservation framework using integration of blockchain and federated learning, SN Comput. Sci. 4 (6) (2023) 703.

[206] D. Alahakoon, R. Nawaratne, Y. Xu, D. De Silva, U. Sivarajah, B. Gupta, Self-building artificial intelligence and machine learning to empower big data analytics in smart cities, Inf. Syst. Front. (2023) 1–20.

[207] M. Faraji-Mehmandar, S. Jabbehdari, H.H.S. Javadi, A self-learning approach for proactive resource and service provisioning in fog environment, J. Supercomput. 78 (15) (2022) 16997–17026.

[208] M. Capra, R. Peloso, G. Masera, M. Ruo Roch, M. Martina, Edge computing: A survey on the hardware requirements in the internet of things world, Future Internet 11 (4) (2019) 100.

[209] S. Iqbal, A.N. Qureshi, K. Aurangzeb, M. Alhussein, S.I. Haider, I. Rida, AMIAC: adaptive medical image analyzes and classification, a robust self-learning framework, Neural Comput. Appl. (2023) 1–29.

[210] P. Zhang, T. Li, G. Wang, D. Wang, P. Lai, F. Zhang, A multi-source information fusion model for outlier detection, Inf. Fusion 93 (2023) 192–208.

[211] B. Marsh, A.H. Sadka, H. Bahai, A critical review of deep learning-based multi-sensor fusion techniques, Sensors 22 (23) (2022) 9364.

[212] T. Li, L. Yang, K. Li, J. Zhai, An intelligent fusion algorithm and its application based on subgroup migration and adaptive boosting, Symmetry 13 (4) (2021) 569.

[213] L. Wu, L. Zhang, L. Chen, J. Shi, J. Wan, et al., A lightweight and multisource information fusion method for real-time monitoring of lump coal on mining conveyor belts, Int. J. Intell. Syst. 2023 (2023).

[214] L. Lei, Y. Zhou, G. Yang, Multisource information fusion-based environment perception and dynamic model of underwater vehicle in irregular ocean environment, Inf. Fusion 94 (2023) 257–271.

[215] N.R. Moloisane, R. Malekian, D.C. Bogatinoska, Wireless machine-to-machine communication for intelligent transportation systems: Internet of vehicles and vehicle to grid, in: 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO, IEEE, 2017, pp. 411–415.

[216] F. Alanazi, M. Alenezi, A framework for integrating intelligent transportation systems with smart city infrastructure, J. Infrastruct. Policy Dev. 8 (5) (2024) 3558.

[217] M. Ren, N. Chen, H. Qiu, Human-machine collaborative decision-making: An evolutionary roadmap based on cognitive intelligence, Int. J. Soc. Robot. 15 (7) (2023) 1101–1114.

[218] B. Geng, P.K. Varshney, Human-machine collaboration for smart decision making: current trends and future opportunities, in: 2022 IEEE 8th International Conference on Collaboration and Internet Computing, CIC, IEEE, 2022, pp. 61–67.

[219] G.C. Saha, S. Kumar, A. Kumar, H. Saha, T. Lakshmi, N. Bhat, Human-AI collaboration: Exploring interfaces for interactive machine learning, Tuijin Jishu/J. Propuls. Technol. 44 (2) (2023) 2023.

[220] F. Zeller, L. Dwyer, Systems of collaboration: challenges and solutions for interdisciplinary research in AI and social robotics, Discov. Artif. Intell. 2 (1) (2022) 12.

[221] P. Bisconti, D. Orsitto, F. Fedorczyk, F. Brau, M. Capasso, L. De Marinis, H. Eken, F. Merenda, M. Forti, M. Pacini, et al., Maximizing team synergy in AI-related interdisciplinary groups: an interdisciplinary-by-design iterative methodology, AI Soc. 38 (4) (2023) 1443–1452.

[222] M. Pawłowski, A. Wróblewska, S. Sysko-Romańczuk, Effective techniques for multimodal data fusion: A comparative analysis, Sensors 23 (5) (2023) 2381.

[223] A. Rawal, J. McCoy, A. Raglin, D.B. Rawat, B.M. Sadler, Recent advances in multi-modal data fusion: status, challenges and perspectives, Artif. Intell. Mach. Learn. Multi-Domain Oper. Appl. IV 12113 (2022) 413–422.

[224] R. Alabduljabbar, H. Almazrou, A. Aldawod, Context-aware news recommendation system: Incorporating contextual information and collaborative filtering techniques, Int. J. Comput. Intell. Syst. 16 (1) (2023) 137.

[225] B. Li, D. Pi, Y. Lin, I.A. Khan, L. Cui, Multi-source information fusion based heterogeneous network embedding, Inform. Sci. 534 (2020) 53–71.

[226] H. Rosmanith, J. Volkert, R. Valles, F. Serrano, M. Plociennik, M. Owsiak, Interactive fusion simulation and visualisation on the grid, in: Sixth International Symposium on Parallel and Distributed Computing, ISPDC'07, IEEE, 2007, 20–20.
[227] F. Yang, P. Zhang, MSIF: Multi-source information fusion based on information sets, J. Intell. Fuzzy Systems 44 (3) (2023) 4103–4112.
[228] I. Mountasser, B. Ouhbi, F. Hdioud, B. Frikh, Semantic-based big data integration framework using scalable distributed ontology matching strategy, Distrib. Parallel Databases 39 (2021) 891–937.
[229] M. Kawas, B. Alkhatib, M. Dashash, Enhancing ontology integration in medical texts through advanced mechanisms, 2023.
[230] M. Dehghan, E. Khosravian, Z. Golfar, H. Shahbazi, P2DF: Privacy-preserving data fusion protocol, in: 2022 12th International Conference on Computer and Knowledge Engineering, ICCKE, IEEE, 2022, pp. 211–218.
[231] Z. Chen, L. Zhao, Q. Li, X. Song, J. Zhang, Multimodal data fusion, in: Advances in Computing, Informatics, Networking and Cybersecurity: A Book Honoring Professor Mohammad S. Obaidat's Significant Scientific Contributions, Springer, 2022, pp. 53–91.
[232] S.E. Hammel, H. Johnson, T. Huber, P. Trenorden, Interactive information preprocessing and control for intelligent data fusion and contact tracking, in: Proceeding of 1st Australian Data Fusion Symposium, IEEE, 1996, pp. 59–64.
[233] P. Prajwal, D. Prajwal, D. Harish, R. Gajanana, B. Jayasri, S. Lokesh, Object detection in self driving cars using deep learning, in: 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, ICSES, IEEE, 2021, pp. 1–7.
[234] L. Tian, D. Turjeman, S. Levy, Privacy preserving data fusion, 2023, Available at SSRN 4451656.
[235] P. Tiwari, S. Viswanath, G. Lee, A. Madabhushi, Multi-modal data fusion schemes for integrated classification of imaging and non-imaging biomedical data, in: 2011 IEEE International Symposium on Biomedical Imaging: From Nano To Macro, IEEE, 2011, pp. 165–168.
[236] X. Shi, X. Sun, M. Fu, Y. Chen, W. Xie, Federated filter for multi-sensor data fusion of dynamic positioning ship, in: 2012 IEEE International Conference on Automation and Logistics, IEEE, 2012, pp. 13–18.
[237] F.P. Martins, J.A. Paixão, C.M. de Farias, F.C. Delicato, Hercules: A context-aware multiple application and multisensor data fusion algorithm, in: 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, DASC/PiCom/CBDCom/CyberSciTech, IEEE, 2021, pp. 197–200.
[238] A. Casanova, A. Pujol-Miró, J. Ruiz-Hidalgo, J.R. Casas, Interactive registration method for 3D data fusion, in: 2016 International Conference on 3D Imaging, IC3D, IEEE, 2016, pp. 1–8.
[239] B. Murray, M.A. Islam, A.J. Pinar, T.C. Havens, D.T. Anderson, G. Scott, Explainable ai for understanding decisions and data-driven optimization of the choquet integral, in: 2018 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE, IEEE, 2018, pp. 1–8.
[240] R.C. Luo, M.-H. Lin, R.S. Scherp, Dynamic multi-sensor data fusion system for intelligent robots, IEEE J. Robot. Autom. 4 (4) (1988) 386–396.
[241] K. Saranya, M. Hema, S. Chandramathi, Data fusion in ontology based data integration, in: International Conference on Information Communication and Embedded Systems, ICICES2014, IEEE, 2014, pp. 1–6.
[242] A. Ligocki, A. Jelínek, L. Žalud, Atlas fusion-modern framework for autonomous agent sensor data fusion, in: 2022 ELEKTRO, ELEKTRO, IEEE, 2022, pp. 1–6.
[243] Z. Lei, P. Cui, Y. Huang, Multi-platform and multi-sensor data fusion based on ds evidence theory, in: 2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology, CCET, IEEE, 2020, pp. 6–9.
[244] F. Zhao, C. Zhang, B. Geng, Deep multimodal data fusion, ACM Comput. Surv. 56 (9) (2024) 1–36.
[245] L. Bernasconi, S. Şen, L. Angerame, A.P. Balyegisawa, D. Hong Yew Hui, M. Hotter, C.Y. Hsu, T. Ito, F. Jörger, W. Krassnitzer, et al., Legal and ethical framework for global health information and biospecimen exchange-an international perspective, BMC Med. Ethics 21 (2020) 1–8.
[246] S. Glasdam, K.Ó. Cathaoir, S. Stjernswärd, Balancing different legal and ethical requirements in the construction of informed consents in qualitative international collaborative research across continents-reflections from a scandinavian perspective, J. Acad. Ethics (2024) 1–14.