A Review of Deep Multi-Objective Reinforcement Learning and Vision-Based Systems for Smart Cities

Saad Alharbi
Taibah University, Janadah Bin Umayyah Road, Tayba, Madinah 42353, Saudi Arabija
E-mail: stharbi@taibahu.edu.sa

Keywords: artificial intelligence, deep learning, computer vision, machine learning, multi-objective optimization, reinforcement learning, smart city

Received: June 18, 2025

Smart cities leverage artificial intelligence to address urban challenges such as traffic congestion, environmental sustainability, public safety, and energy efficiency. Among AI techniques, the integration of multi-objective reinforcement learning (MORL) and computer vision (CV) offers adaptive, real-time decision-making capabilities while processing complex visual data streams. This paper presents a comprehensive review of the joint application of MORL and CV in smart city environments. A systematic search was conducted across six major databases (Scopus, Web of Science, IEEE, Springer, MDPI, and Elsevier) from 2019 to 2024, resulting in the selection of 90 relevant studies. The review follows a thematic analysis approach, categorizing the literature into smart mobility, infrastructure, environment, governance, and smart living. Key findings indicate that multi-agent MORL and CV are increasingly used in traffic signal control, autonomous vehicle navigation, energy management, surveillance, healthcare, and waste management. However, despite advancements in deep RL algorithms like DDPG, PPO, SAC, and advanced CV techniques such as semantic segmentation and multi-camera tracking, the direct integration of these technologies remains underexplored in many domains. The paper highlights current research gaps, including the lack of standardized frameworks for MORL-CV synergy, scalability limitations, ethical concerns, and insufficient quantitative benchmarking across studies. Additionally, trends such as federated learning, edge computing, and digital twins are identified as promising enablers for future MORL-CV solutions in urban contexts. This review serves as a resource for researchers and policymakers aiming to develop sustainable and intelligent urban systems by bridging perception (via CV) with adaptive control (via MORL) for real-time, multi-criteria decision support.

Povzetek: Prispevek se ukvarja s povezovanjem večciljne okrepitevne poti (MORL) in računalniškega vida v pametnih mestih. S sistematično analizo 90 študij pokaže rastočo uporabo MORL–CV v prometu, energiji, nadzoru in zdravstvu in izpostavi vrzeli: odsotnost standardiziranih okvirjev, šibko neposredno integracijo ter omejitve glede razširljivosti in etike.

1 Introduction

In 2018, 55% of the population globally lived in urban areas, and the statistic is estimated to be 68% by 2050, according to the United Nations Department of Economic and Social Affairs [1]. This significant population shift toward cities has created major urban challenges, including traffic congestion, pollution, safety issues, and inefficient resource consumption. These complex and dynamic problems demand intelligent and adaptive solutions, giving rise to the concept of smart cities. For instance, megacities like Mumbai and Jakarta face severe congestion and pollution due to outdated infrastructure and rapid growth. A smart city employs technologies like Artificial Intelligence (AI), the Internet of Things (IoT), and big data to optimize infrastructure and urban services.

Among these, reinforcement learning (RL) and computer vision (CV) are key technologies that can provide efficient and scalable solutions to various urban issues. RL can assist in making optimal decisions based on environmental feedback, while CV helps machines interpret and understand visual inputs, enabling real-time monitoring and analysis. The integration of RL and CV in smart city systems has gained attention in recent years. RL, particularly multi-objective reinforcement learning (MORL), is effective in addressing urban systems' conflicting goals, such as reducing congestion while minimizing energy consumption. MORL techniques such as scalarization, Pareto-based optimization, and decomposition methods are particularly useful in balancing these competing objectives. Simultaneously, CV contributes through advanced techniques in image classification, object detection, and motion tracking that are highly relevant in urban scenarios like traffic analysis, surveillance, and waste monitoring. Recent techniques such as semantic segmentation and vision transformers have further enhanced scene understanding in smart environments.

Recent advances in deep learning have enhanced the effectiveness of both RL and CV in urban applications.

For example, deep Q-networks (DQN) and proximal policy optimization (PPO) have enabled scalable RL, while CNNs and transformers improved object recognition accuracy. Despite the potential and growing interest in these technologies, research that comprehensively examines the synergy between RL and CV in smart city applications remains limited. Existing studies often focus on either RL or CV in isolation, or they address specific application areas without offering a holistic view. While surveys exist on RL in traffic optimization [2] and CV in urban monitoring [3] , few examine their joint application across smart city systems. Figure 1 illustrates major technological milestones in RL and CV that have shaped their application in smart cities.

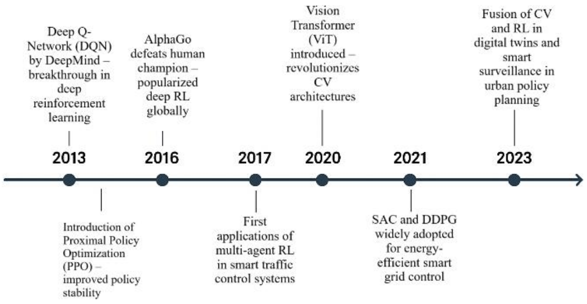

Figure 1: Major technological milestones in RL and CV

A focused review that explores the overlap between these fields in the context of smart cities is needed to highlight current advancements, identify research gaps,

This paper aims to fill this gap by reviewing the literature on multi-objective reinforcement learning and computer vision, specifically focusing on their combined use in smart city contexts. It categorizes and analyzes recent studies, summarizes trends and methodologies, and outlines the key challenges and open research questions in this emerging domain. Figure 2 illustrates the conceptual framework underpinning this review. It highlights how smart city applications leverage data sources (e.g., IoT devices and sensors) that are processed by Computer Vision (CV) for perception tasks and by Reinforcement Learning (RL) for adaptive decision-making and policy optimization. This integration enables real-time, multi-objective optimization across key urban domains such as mobility, infrastructure, governance, environment, and living.

Recent breakthroughs have accelerated the synergy between RL and CV for urban systems. Key milestones include DeepMind's AlphaGo (2016), which demonstrated the power of deep RL for complex decision-making [4], and the rise of Vision Transformers [5], which revolutionized visual feature extraction and scene understanding. These advancements, along with developments in federated learning and edge AI [6], underpin many of the smart city applications reviewed in this paper.

To address this gap, the following research questions guide this review:

RQ1: How are Deep Multi-Objective Reinforcement Learning (MORL) and Computer Vision (CV) currently integrated in smart city applications?
RQ2: What are the dominant algorithmic strategies, challenges, and limitations in applying MORL and CV jointly for urban decision-making?
RQ3: What are the emerging trends, gaps, and future research directions for enabling real-time, multi-criteria decision support systems combining MORL and CV?

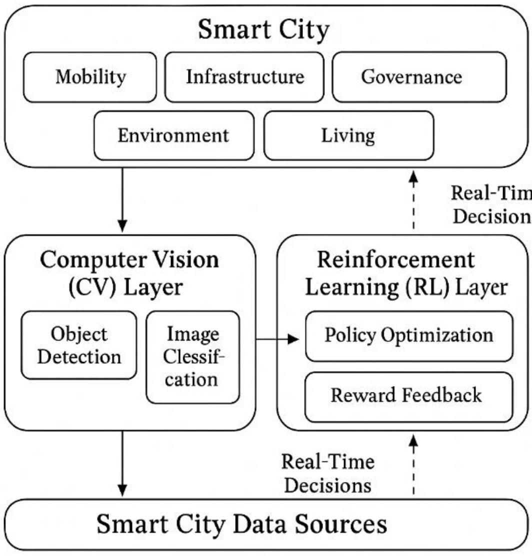

Figure 2: Conceptual Framework of RL and CV Integration in Smart City Systems (Source: Author)

2 Method

2.1 Search strategy

The current review gathered data from published articles from peer-reviewed journals and renowned databases like Scopus, Web of Science, IEEE, Springer, etc. (See Table 1). These databases were selected due to their high relevance, coverage of top-tier AI/ML journals, and indexing of conference proceedings in the smart city domain. Data was searched based on investigating the role of ML and multi-objective RL specific to CV in enhancing smart city development. The search strategy employed the following Boolean string to maximize retrieval of relevant studies: ("multi-objective reinforcement learning" OR "MORL" OR "deep reinforcement learning" OR "DRL") AND ("computer vision" OR "vision-based systems") AND ("smart cities" OR "urban systems"). Search queries were customized for each database, using relevant syntax and filters to exclude non-related computational studies and retain only smart city applications. The review period was set between 2019 and 2024, and only English, peer-reviewed, full-text articles were considered.

This pairing enables smart city systems to process visual input (via CV) and make adaptive policy decisions (via RL), a combination increasingly required in dynamic urban environments. This search gathered data from 90 literature studies from the last six years (2019–2024) to collect the most recent advancements in ML-based developments in smart cities. Initial screening involved evaluating article titles and abstracts for relevance, followed by full-text review for methodological alignment with smart city applications. Studies were selected specifically for the keywords: Artificial Intelligence (AI), Machine Learning (ML), Multi-objective Reinforcement Learning (RL), Deep Learning (DL), Computer Vision (CV), Smart Cities, Decision Making, Smart Mobility, Smart Infrastructure, Smart Governance, Smart Economy, Smart Environment, Smart Living and Smart People. Synonyms and related terms such as “urban computing,” “perception systems,” or “adaptive control” were also considered where applicable. Search keywords were combined with Boolean operators like AND, OR, WITHIN, etc., to create a relevant search strategy and boost data collection search. The query syntax and filtering parameters were customized per database to maximize relevance and exclude unrelated computational studies.

Table 1: Data selection strategy

Years	Search Engines	Keywords
2019-2024	Google Scholar Scopus Web of Science IEEE Springer MDPI Elsevier Wiley	Artificial Intelligence (AI) Machine Learning (ML) Multi-objective Reinforcement Learning (RL) Deep Learning (DL) Computer Vision (CV) Smart Cities Decision Making Smart Mobility Smart Infrastructure Smart Governance Smart Economy Smart Environment Smart Living Smart People

The selection process followed a PRISMA-inspired flow:

- Initial records identified: 312
- After duplicate removal: 280
- Full-text assessed for eligibility: 125
- Studies included in the final review: 90
- Figure 3 illustrates the selection process.

2.2 Inclusion and exclusion criteria

The articles were filtered based on criteria set for inclusion and exclusion based on the study’s relevance. The term "relevance" here specifically refers to the use of reinforcement learning (RL) and computer vision (CV) in smart city contexts. Table 2 was used as instrument in narrowing the broader set of retrieved articles to the 90 selected for this review, ensuring methodological and topical consistency.

2.3 Data extraction

Data extraction followed the six-step thematic analysis process proposed by **Braun & Clarke (2006)**, including: (1) familiarization with the data, (2) initial code generation, (3) theme identification, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report [7]. Such approach commonly applied in structured ML reviews such as [8]. This process was adapted to the technical context of RL and CV applications in smart cities. This process followed the widely recognized methodology involving familiarization, coding, theme development, and refinement, adapted to suit the technical focus of the current study. Each article selected for review was explored based on the key outcomes, impacts on smart city applications, and concerned limitations. Key outcomes included aspects such as performance metrics, scalability, and practical deployment challenges in smart city environments. Repetitive keywords were used in the initial coding of articles to ensure robustness. Excel spreadsheets were used allowing categorization by recurring concepts such as optimization objectives, visual analytics, and policy learning mechanisms. Later, based on data coded, themes were developed, including: “smart city domains”, “Multi-objective RL and CV in smart cities for intelligent decision making”, and “applications of RL in CV in smart cities”. These themes were foundational in organizing the review sections and synthesizing cross-domain findings systematically.

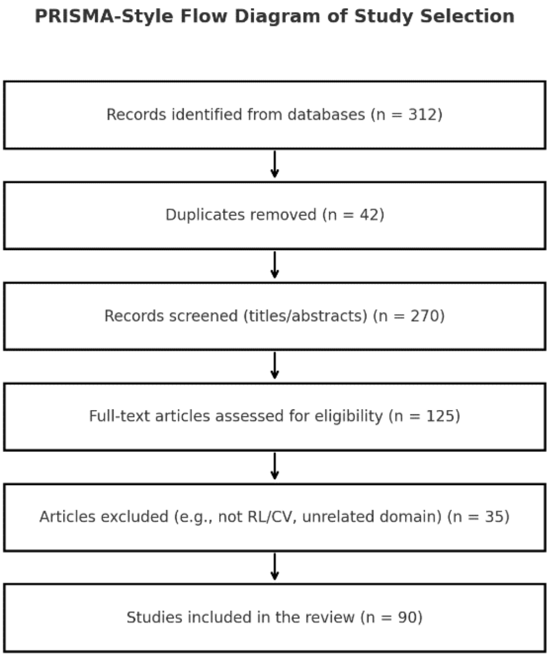

Figure 3: PRISMA-style flow diagram of the literature selection process adapted for AI/ML-based smart city review (Adapted from PRISMA 2020 guidelines [9]).

The articles were coded into categories such as:

- Smart city domain (e.g., mobility, governance, infrastructure)

- Algorithmic approach (e.g., MORL, CV, joint RL-CV integration)
- Performance metrics (e.g., accuracy, energy savings, latency)
- Deployment setting (simulation or real-world)
- Data extraction was performed using Excel spreadsheets to ensure consistency and replicability.

Table 2: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
All the studies included were published in English in peer-reviewed journals in full-text format.	Studies not published in peer-reviewed journals were excluded.
Studies published between 2019 and 2024 were included.	Studies published before 2019 were excluded.
Studies exploring the implications of ML, specifically RL and CV, in smart cities were included.	Studies discussing smart cities otherwise were excluded.
Quantitative, qualitative, and mixed-methods studies were included.	None were excluded based on their methodology.

2.4 Glossary of acronyms

This section provides a glossary of key acronyms and technical terms used throughout the paper.

- **RL** – Reinforcement Learning, a machine learning paradigm for sequential decision-making.
- **MORL** – Multi-Objective Reinforcement Learning, an RL approach handling multiple (possibly conflicting) objectives simultaneously.
- **CV** – Computer Vision, a field of AI that enables machines to interpret and process visual data.
- **DRL** – Deep Reinforcement Learning, an RL approach utilizing deep neural networks for complex policy learning.
- **DDPG** – Deep Deterministic Policy Gradient, a model-free RL algorithm for continuous action spaces.
- **PPO** – Proximal Policy Optimization, a stable policy-gradient-based RL method balancing exploration and exploitation.
- **SAC** – Soft Actor-Critic, an off-policy RL algorithm that improves convergence speed through entropy maximization.
- **MA2C** – Multi-Agent Actor-Critic, used for decentralized policy learning in multi-vehicle systems.
- **MOMCT** – Multi-Objective Multi-Camera Tracking, a CV technique for cross-camera object tracking and re-identification.
- **GAN** – Generative Adversarial Network, a deep learning model composed of generator and discriminator networks trained adversarially.
- **CNN** – Convolutional Neural Network, a deep learning architecture designed for image-based tasks.
- **ViT** – Vision Transformer, a transformer-based model for image classification using attention mechanisms.

3 Results

This section presents the findings of the reviewed studies, structured across five key smart city domains: mobility, infrastructure, governance, environment, and living. These domains were identified through thematic analysis, aligning with the smart city framework. Each subsection highlights the reinforcement learning (RL) and computer vision (CV) methods applied, their degree of integration, and domain-specific applications. Where available, performance metrics such as accuracy, scalability, and deployment challenges are discussed. This structure facilitates not only a comprehensive synthesis of state-of-the-art methods but also a critical identification of cross-domain trends, technical contributions, and existing research gaps.

Across the 90 reviewed studies, CV-based models for traffic detection and surveillance achieved an average detection accuracy of 82–88% mAP, with multi-camera tracking systems (e.g., MOMCT) reporting IDF1 scores above 80%**.** RL-driven energy management solutions demonstrated 8–20% energy savings compared to static scheduling baselines, while MORL approaches for autonomous mobility (e.g., MA2C and PPO) achieved 15–20% reductions in CO_2 emissions in simulation benchmarks. These metrics highlight the growing maturity of RL-CV pipelines in mobility and energy domains, though real-world deployments remain limited.

3.1 Smart city domains and CV

CV in many smart city domains enhances the safety, efficiency and security of smart cities by integrating efficient ML models. This integration supports critical tasks such as real-time surveillance, resource optimization, and autonomous control systems. The purpose of combining ML with CV is to advance edge detection, semantic segmentation, object detection, object recognition, text detection and others in various applications and spheres of smart cities. These functions form the perceptual layer of smart systems and are pivotal for interpreting urban sensory data in context-aware environments. These specific domains include smart mobility, smart governance, smart education, smart healthcare, smart economy, smart governance, smart living, smart environment, smart people & infrastructure [1], [10], [11] [12]. Table 3 shows the significant applications of all these smart city domains.

Smart mobility, connected with the goals of smart city expansion, solves the problem of traffic congestion and the environmental effects in large cities. It aims to ease the mobility of individuals and goods in a town, generating six benefits: traffic, transport duration, the cost incurred in transport, pollution, noise and safety experienced when using the transport means. It fosters connected, sustainable and participative mobility, leveraging IoT, AI, Blockchain, and Big Data (BD) technologies to design and develop sustainable and disruptive solutions that rewrite the paradigm of cities and their inhabitants [13], [14], [15]. Smart mobility is essential in the enhancement of smart cities across the globe. In this case, double-parking and other roadside activities that are close to each other hinder traffic in areas with high transport density. The real-time IoT-based CVROSS, namely computer vision-based roadside occupation surveillance system, employs smart HD cameras and wireless modules. It watches traffic images dynamically, considers whether people are beside the road, and enhances image quality. Developed and piloted in Hong Kong, CVROSS is targeted to improve traffic and fleet for smart mobility [16].

Intelligent transport systems (ITS) provide solutions for smart infrastructure, mobility, traffic management and safety, and efficient, effective and secure means of transport [17], [18]. They are anticipated in areas such as public transport management, smart structure management, intelligent route management, superior vehicle control, integrated bill payment, and route details. Components such as global positioning system tracking aids on transit, technology in sensing, and video analytics assist in alleviating pollution and optimizing resources [19]. Smart transportation deals with safety, efficiency, and security in government rules, vehicle support services, and transportation payment services [10]. Therefore, the authors have put forward a procedure used for estimation of the wind turbine blades' angular velocity in a real wind farm scenario using ML and CV approaches. They use a WSN to capture videos and send them to a coordinator that computes the enhanced CV and ML Algorithms. The model can also be used to gather and parse more data about wind turbines such as mechanical distortions, surface erosions and super heating points in different wind turbine sections. Intelligent Energy Management Systems (IEMS), therefore, seek to optimize efficiency in electrical power generation, flexibility, renewable systems, and low carbon emanation through Machine Learning which predicts efficiency with IoE network support [13], [20], [21]. Smart education is associated with a technological enhancement in the classroom environment that brings about a smart classroom to facilitate tele-education solution similar to the physical classroom with voice recognition features using CV. This has shifted the use of libraries from places of study and quiet to discourse, peer, social learning and an exhibition hall. Lately, some institutions of higher learning are embracing technology to make their universities and libraries smart spaces as part of the new education trend; such advancements boost the competitiveness of a smart city.

Table 3: Smart city domains and involvements

Smart City Domain	Description	Involvements
Smart Mobility	It offers smart and intelligence mobility and transportation.	Traffic Monitoring and Management Smart Supply Chain Management Smart Parking and Routing Sustainable and Autonomous Mobility
Smart Living, Smart People & Infrastructure	It offers advancement of the city's infrastructure while improving the quality of life.	Smart Communication Smart Energy and Grid Smart Home Smart Community
Smart Education	It enhances the need to meet the emerging digital native generations' demands.	Virtual Reality based learning Student Management Smart Library Smart Classroom
Smart Governance	An application of innovation and technology to improve planning and decision-making in governing organizations.	Smart Urban Management Smart Urban Planning Smart Building Disaster Prevention and Management E-governance
Smart Economy	It is centered on technical innovation, resource efficiency, sustainability, and high social wellbeing.	Smart Supply Chain Smart Commerce E-commerce Retail
Smart Environment	An application of creating an environment with integrated sensors for a better environment.	Waste Management Smart Irrigation Smart Weather Monitoring Air Quality Monitoring
Smart Healthcare	A method for healthcare delivery uses IoTs, wearable tech, and mobile internet for enhanced resources, connecting people, data and institutions.	Telemedicine Smart Hospital E-health record Telenursing Patient Monitoring

The significance of IoT, ML, and CV for a smart education system lies in the use cases in a smart electric learning situation, smart classrooms, smart libraries, and smart attendance systems [22], [23], [24] .

Smart healthcare integrates health monitoring systems to determine the mobility of the human body using vision detection or sensors [25]. They are used in identifying unusual behavior among patients and help prevent avoidable deaths due to health complications. Patient monitoring systems will be useful in the next five to ten years, and vision- and sensor-based detection techniques will be useful [10]. Additionally, studies examined that telemedicine makes it possible and efficient for specific groups of patients: long-term care patients, specialized demand patients, and time-saving patients, and those patients who cannot overcome financial/ geographical barriers. It is in this capacity that disease management programs, with regards to the online space, enhances the delivery of healthcare services, resources, and patient care [26], [27], [28]. The telemedicine of the digital age that offers interactive discussions and video-conferences enhances the standard healthcare, the research data, and medication dosages. It facilitates treatment for medical practitioners from different geographical areas, whether in health centers or rural or urban areas [29], [30]. Overall, CV-based solutions for traffic and surveillance demonstrate higher performance maturity (e.g., detection accuracy often above 80% mAP in large datasets), whereas applications like smart healthcare and education remain experimental and lack standardized benchmarks. The absence of large-scale, multi-modal datasets integrating visual data with RL pipelines is a recurring gap highlighted in recent studies [21]. These CV tasks, when combined with RL-based decision-making policies, enable adaptive optimization. For example, outputs from multi-camera object detection can serve as state inputs for RL traffic controllers, enabling real-time congestion management.

3.2 RL in smart cities for intelligent decision making

RL is a type of ML that learns through trial-and-error interactions with the environment, aiming to optimize cumulative reward. In the context of smart cities, it enables systems to learn policies that adapt to dynamic urban environments, supporting real-time decision-making and autonomous control. Its applicability in smart cities revolves around policy learning and optimization, essential for autonomous systems that operate under uncertain and changing conditions. In various reviewed studies, RL were implemented in applications such as traffic signal control, resource management, and energy consumption optimization. Such implementations demonstrate RL's potential to replace static rule-based systems with adaptive, feedback-driven agents.

Multi-objective routing optimization solves barriers of ITS efficiently and rapidly when modelling uncertainty of time for transportation, resolving challenges of carbon emission cost, transportation cost and transport time using deep Q-learning algorithms like Artificial Fish Swarm Optimization (AFO), GA, Particle Swarm Optimization (PSO). DDPG is an approach defined by deep Q-learning with a continuous action space for physical tasks in simulating ITS [31], [32]. Two new approaches to multi-agent power control in device-to-device (D2D) communications based on a 5G network vehicle-to-vehicle (V2V) communications using the DDPG algorithm [33]. Compared with other deep RL methods, the developed models have higher network energy efficiency and flexibility for smart city applications and high-quality video streaming in real-time. To synthesize the differences between the commonly applied DRL algorithms, Table 4 compares DDPG, PPO, and SAC in terms of strengths, limitations, smart city use cases, and reported performance outcomes. Another work also presented a two-timescale federated deep RL framework named FDPG based on the DDPG in the smart city scenario to deal with the combined problem of resource allocation and task offloading under the premise of privacy preservation. This approach is intended to reduce the energy consumption of IoT devices under the given delay constraint [34], [35]. Some of the classification methods are CV-based solutions that effectively segregate the object from the heap of garbage and trash from the various waste objects that are commonly available, such as paper, paper boxes, food, glass, etc. Since the DL and Deep RL (DRL) techniques have recently advanced, the waste identification and detection make waste object classification possible. Thus, an intelligent DRL-based recycling waste object detection and classification (IDRL-RWODC) model for smart cities is proposed in this aspect. The IDRLRWODC technique should ideally aim at accurately identifying and categorizing waste objects, which can be achieved using the DL and DRL methods. The IDRL-RWODC model involves two processes: Mask Regional Convolutional Neural Network (Mask RCNN) based object detection and DRL-based object classification. Furthermore, the DenseNet model is used as a baseline model for the proposed Mask RCNN model. As a classifier, researchers used a deep Q-learning network (DQLN) [36]. PPO is one of the most common multi-objective RL algorithms [37] ; it is an advantage that alternates between exploitation and exploration, which benefits smart cities [3]. It controls the traffic signals depending on the data of the IoT sensors to decrease traffic and increase the ecologically friendly environment in urban areas. It can bring adaptive solutions for energy management and resource distribution according to current consumption rates. Therefore, using smart surveillance systems, PPO can increase public safety and detect disturbances at early stages when responding to potential threats [38], [39], [40]. Thus, PPO depends on data collection and analysis and the subsequent improvement in its functions that contribute to developing efficient and sustainable urban areas. PPO is applied to navigate autonomous vehicles using DRL and was therefore applied to this circumstance. It has been used to regulate the choice of lanes during the changeover and mixed traffic signals to guarantee vehicle safety at the interphase [41], [42].

A relatively new multi-objective DRL algorithm known as the soft actor-critic (SAC) is implemented to derive the best solution for energy distribution in EVs with battery-supercapacitor HESS. The recommended SAC-based EMS improves drawbacks in the current DRL techniques, including lower convergence speed and erratic training behavior [43]. The algorithm is embedded in the algorithm is embedded in an environment with continuous actions through self-play along with a newly designed reward signal. The results also confirm that the designed SAC-based EMS outperforms rule-based methods, deep deterministic policy gradient, and battery-centric setup. Using SAC algorithm, the proposed MASAC model for training arterial traffic control is superior to conventional MARL algorithms and multiband-based strategies [31], [44], [45].

Past studies have shown that innovative Reinforcement Convolutional Transfer Learning (RC-TL), a trajectory-prediction system based on CNN solves multi-objective and discrete problems. This system groups users with such backgrounds; each group has an RL agent to improve a CNN neural design; one model is trained per cluster from a limited user sample; the model is then shared with other users in the cluster. It is also helpful for object recognition of vehicles and pedestrians, automatic parking, autonomous driving, handover and service migration, network control, and caching such services in edge-computing and mobile networks [44], [46].

Multi-objective RL neural networks are also applied to smart cities for smart energy control [47]. The model-free control methods like RL enable easier power deployment without having to derive or find a thermal model of the building. HVAC, DHW, and home lighting systems have been controlled using artificial neural networks (ANN) to decrease energy bills and discomfort. RL algorithms have been shown to provide near-optimal solutions for maintaining energy in low-exergy buildings and multi-objective deep RL controllers have been developed to reduce electricity consumption and solve problems of energy resource management [45], [48]. However, most of the contributions incorporate simplified building energy models or specific RL applications and, therefore, pose the need to expand the high-performance simulation model with enhanced ML techniques [49], [50], [51], [52].

There has been continued refinement of unmanned automated vehicles (UAVs) and progressive control systems, with the concept of smart cities targeting self-driving automobiles for task offloading and trajectory control. A multi-objective algorithm for car mobility in road intersections is designed using a hierarchy reinforcement learning (HRL) method. The specific objective of comparing the performance of a signalized vehicular intersection with a fixed-time traffic controller using HRL can be as efficient as traditional traffic controllers [2], [53]. Self-driving vehicles require intelligence to make sound decisions and actions in traffic situations. However, many cases of single optimization/sampling-based motion planners and end-to-end learning methods fail to produce safe trajectories in real-time. A hierarchical behavior planning architecture, which consists of low-level safe controllers and one high-level DRL (H-CtRL), is thus proposed. Low-level controllers ensure safety, while the H-CtRL is adaptive and efficient. This algorithm was tested with a simulator that eliminated the chance of failing in real-life situations, as was tested and proven successful in smart cities [41], [54].

To provide a structured comparison of existing approaches, Table 5 presents a summary of recent studies combining Multi-Objective Reinforcement Learning and Computer Vision techniques in smart city applications. The comparison outlines the algorithmic strategies, application domains, vision tasks, optimization objectives, and reported performance metrics. Additionally, the table highlights the degree of integration between RL and CV in each study, distinguishing between independent use, parallel deployment, and fully integrated pipelines. This dual-purpose analysis addresses both the literature synthesis gap in related works and the need for cross-domain comparative insight identified in the current review findings.

Comparative studies show that PPO and SAC achieve faster convergence and better adaptability for dynamic traffic and energy systems, while DDPG excels in continuous control but struggles with sample inefficiency [55] [56]. Despite their potential, most RL solutions remain evaluated in simulations, with few real-world deployments at city scale.

The DRL algorithms discussed (DDPG, PPO, SAC) often rely on perception modules powered by CV (e.g., object detection or motion tracking) to build dynamic environment states. This fusion of perception and control forms the foundation for intelligent decision-making pipelines in smart cities.

3.3 Applications of multi-objective RL and CV in smart cities

3.3.1 Smart mobility

For sustainable urban mobility, multi-objective RL and CV advance smart mobility management, enhancing shared mobility and public transport, advancing public transport optimization, traffic management, autonomous vehicles, and smart parking [56], [61], [62], [63] [64]. These techniques enable learning-based adaptation to real-time traffic states, making them suitable for dynamic routing, congestion pricing, and policy simulations.

An example of the presented multi-objective algorithms is a Multi-Agent Actor-Critic (MA2C) for multi-AV lane-changing in a mixed traffic environment, which is one of the significant parts of ITS in smart cities. RL is then used to optimize the behavior of the algorithm with respect to lane-changing, leading to better environmental impact and quality of life standards for occupants of urban areas. It also provides a realistic imitation of driver-controlled vehicular dynamics that are useful in traffic flow control. By addressing lane change control, passenger comfort, and inter-vehicle cooperation, the efficiency of urban traffic flow is maintained without

Table 4: Comparative Overview of Key DRL Algorithms in Smart Cities

Algorithm	Strengths	Limitations	Example Use Cases	Performance Highlights
DDPG	Effective in continuous action spaces; suitable for complex control tasks	Requires careful hyperparameter tuning; sample inefficient	Smart energy management and V2V communications [34] [57]	Achieved higher energy efficiency but slower training compared to PPO
PPO	Stable training, robust to hyperparameters; balances exploration and exploitation	Moderate computational cost	Traffic signal optimization, smart surveillance [58], [59]	Reduced congestion and improved traffic flow by 10–15%
SAC	High sample efficiency; faster convergence due to entropy maximization	Sensitive to reward design and tuning	Energy distribution in EV systems[13] [60]	Outperformed rule-based and DDPG methods in energy savings and scalability

Table 5: Summary of RL and CV Integration in Smart City Applications, Comparative overview of algorithms, domains, integration types, and performance metrics across recent studies

Study	Algorithm/Model	Application Domain	CV Task (if any)	RL Objective	Performance Metrics
[61]	MOMCT (Multi-Objective Multi-Camera Tracking)	Smart Mobility	Object Tracking & Re-identification	Optimize multi-camera association	mAP 85.3%, IDF1 80.1%
[56]	MA2C (Multi-Agent Actor-Critic)	Traffic Management	Traffic Flow Detection	Lane changing, emission reduction	Travel time ↓ 15%, CO2 ↓ 12%
[20]	IEMS with RL	Energy Management	Sensor-based Vision	Power grid optimization	Energy savings ↑ 10%
[21]	CV + ML for Wind Turbines	Renewable Energy	Turbine Blade Motion Detection	Angular velocity estimation	Error ↓ 5%, Remote accuracy ↑
[65]	H-RL-VaNSAS	Public Transport	GIS & Route Mapping	Bus route optimization	Distance ↓ 8%, Safety ↑
[66]	PMORL	Healthcare/Smart Living	Patient Activity Recognition	Energy allocation & emissions	Emission ↓ 20%, Accuracy ↑
[60]	SAC-based Traffic Control	Smart Infrastructure	Visual traffic input	Traffic signal optimization	Throughput ↑ 10%, Delay ↓

congestion and less emissions are produced compared to other models [56]. The use of decentralized agents improves scalability and allows for localized learning while preserving global system performance.

Multi-Objective Multiple Camera Tracking (MOMCT) is a critical challenge in the field of computer vision and security involving a search for multiple objects in a scene and tracking their movements from the video shot by various cameras. Distribution learning based on MOMCT in intelligent transportation, public security and self-driving introduces the main object detectors, discusses the advanced level of evaluation information and summarizes benchmark datasets [61]. MOMCT is also essential for cross-camera identity association and spatiotemporal consistency, both crucial in smart surveillance systems. Combining multi-objective RL and ML constant avoidance of heuristics made by human beings can solve combinatorial optimization issues. These methods are faster than metaheuristic methods, particularly for big problems that are useful in IoT-based smart cities. However, RL and ML applications in combinatorial optimization integrate multiple algorithms, for instance, shuffled frog leaping, graywolf, earthworm, and simulated annealing that have been used on power management, traffic routing, traffic police scheduling, and home energy management. The integration of hybrid metaheuristics with policy-learning-based RL enhances exploration capabilities and accelerates convergence on large-scale problems. Further, the deployment of self-driving cars is considered optimal in urban livery, such as business application Fleet Management System and consumer-use ride-sharing apps [62], [63], [67]. Despite these advancements, current studies largely implement RL for mobility optimization and CV for traffic perception as separate components. Integrating real-time visual inputs—such as vehicle and pedestrian detection—into MORL-based traffic policy adaptation remains an underdeveloped research direction. Among mobility-focused studies, MA2C-based lane optimization shows up to 12–15% CO_2 reduction and better scalability than PPO or traditional MARL approaches. However, few methods leverage CV outputs (e.g., multi-camera tracking) as direct inputs for real-time RL-based traffic signal control, which remains a critical research opportunity.

CV tasks like vehicle and pedestrian detection (e.g., MOMCT) feed real-time visual information to RL-based traffic signal optimization and lane-changing strategies, improving both safety and flow efficiency.

3.3.2 Smart infrastructure

DL algorithms efficiently support smart infrastructure development by proposing high-end applications of smart grids, smart building management, energy-efficient behavior and management optimization, renewable power-based energy storage, and infrastructure monitoring. These applications involve multi-objective problems and optimization algorithms [68], [69]. In these domains, DL models offer real-time processing capabilities that outperform traditional statistical methods, particularly in high-dimensional energy consumption datasets.

Residential energy management using different ML algorithms for Building Energy Management Systems (BEMS) is developed to achieve lower power consumption and energy costs. These are demand-side management schemes, residential load scheduling, hybrid optimization algorithms, hybrid home energy management systems (HEMS), and optimization of home energy management systems incorporated with renewable sources of energy. This has been assessed using the decision trees as well as the Gaussian Naive Bayes and the K-Neighbors, feed-forward multilayer perceptron (MLP) neural network ML and DL models. The comparative performance of these models often varies depending on seasonality, household size, and the granularity of energy consumption data.

RL is used to gain energy conservation significance to climate control, lighting, ventilation, heating, and air conditioning (HVAC) applications that crest to 10% and for water heaters crest of 20% energy conserved for smart infrastructure [69], [70], [71], [72]. Adaptive RL agents can adjust HVAC schedules based on occupant patterns and weather forecasts, enhancing comfort without compromising energy goals. Additionally, the Improved Whale Optimization Algorithm (IWOA) is used to optimize domestic appliance design in the HEMS, which is a smart grid tool [41]. The use of IWOA enables better solution diversity and convergence speed in highly non-linear design spaces, making it suitable for real-time smart grid decision support systems. A critical observation is that smart infrastructure research typically applies MORL for energy management and CV for structural monitoring in parallel but not in synergy. Future work should integrate CV-based occupancy detection into RL-driven adaptive energy management systems. Most RL-based energy management systems achieve 8–20% energy savings compared to static scheduling [20], but almost none integrate real-time CV-based occupancy detection for adaptive reward shaping. Future designs should merge these two pipelines to improve context-awareness and building automation.

While RL dominates energy optimization in infrastructure, integrating CV (e.g., occupancy detection or visual inspection of grids) as a feedback signal can enhance adaptive energy management policies.

3.3.3 Smart governance

Advancements in multi-objective digital twin technology have made governments able to integrate smart governance applications monitoring urban management among people, environment and infrastructure through digital identity systems and civic engagement through intelligent surveillance [73], [74]. It enhances the transparency and efficiency of the administrative processes involving AI, data analytics, and ML to streamline decision-making in urban administration [75]. Digital twins enable simulation-based governance, allowing predictive modeling of city-wide scenarios such as disaster management, energy use, or urban planning interventions.

The application of AI chatbots in smart governance in smart cities has relied on the approaches to rule-based chatbots and natural language processing for customer support. They also reflected on the fact that multi-agent systems also serve the purpose of intelligent governance of administrative procedures by integrating AI and IoT in e-government services. These systems improve scalability and responsiveness in citizen service platforms, particularly when integrated with cloud infrastructure and edge devices. To develop smart cities, improve citizen relations, and offer quality services, AI and ML work efficiently. The optimal approach to identifying the extent of the positive impact of AI in the provision of urban public services is based on the dynamic interaction between urban public services and AI technology [75]. This interaction can be modeled as a feedback loop between user behavior data and adaptive service models, enhancing personalization and performance over time.

Moreover, to support smart governance, big data technology collects multi-source and multi-objective heterogeneous data to enhance the RL for urban governance. The author explains a cooperative traffic signal control approach based on a swarm RL approach, which is a hybridization of PSO and RL. Swarm RL techniques allow decentralized control policies that can adapt in real time to traffic flow variability and urban disruptions. The method outperforms other single RL-based adaptive traffic signal control methods, infrared sensors, Identification Devices (RFIDs), and GPS systems facilitating the collection and evaluation of real-time data [74]. Despite the adoption of intelligent surveillance and multi-agent RL in governance, there is limited exploration of frameworks where CV-based urban observations directly inform adaptive RL-driven policy adjustments.

This represents a key gap in current research. Current governance solutions focus more on RL-based policy optimization than on fusing visual surveillance data into adaptive decision-making loops. This imbalance limits their effectiveness in real-time urban risk detection and disaster management.

Intelligent surveillance systems using CV can provide situational data that multi-agent RL systems use for decentralized governance decisions (e.g., adaptive traffic signal policies during emergencies).

Although CV is less prominent in smart infrastructure compared to RL-based control systems, emerging works explore CV for fault detection and visual monitoring of energy grids. These CV outputs can be fused with RL models to enhance decision-making in smart building automation and predictive maintenance systems.

3.3.4 Smart environment

Smart environment entails proper utilization of the available resources, especially in energy usage using multi-objective RL. Smart grids use AI in order to increase the effectiveness, accuracy and eco-friendliness of the overall system. These new electrical networks incorporate information technology in electricity generation, distribution, and use. The integration of distributed sensors, IoT devices, and data analytics allows smart grids to perform real-time adjustments based on consumption patterns and grid status. Deep RL techniques are employed to find the best way to trade energy and perform demand response mechanisms in a Peer-to-peer (P2P) energy system to reduce costs at the household level. In these systems, RL agents learn optimal trading strategies while considering dynamic pricing, grid constraints, and user preferences, enhancing both individual and collective energy efficiency. In green smart cities and electricity consumption, advanced ML and RL techniques are applied, like the LSTM and RNNs, to enhance the dynamic characteristics of smart grids [75].

Additionally, another study introduced a new short-term electricity load forecasting method based on ML and RL-CNN approaches, which improved accuracy and extended the running time. RL-CNN hybrids combine spatial pattern recognition with adaptive reward optimization, improving both prediction accuracy and the adaptability of forecasting systems to seasonal shifts. A new smart energy management system based on RL was also introduced that uses Markov decision processes and Q-learning to learn energy allocation via real energy trading. The proposed energy management algorithm proves to show the numerical efficacy of reduced electricity costs [74]. Markov models support the modeling of sequential decision-making in environments with probabilistic transitions, making them highly relevant for grid stability under demand uncertainty. While CV technologies monitor environmental indicators such as air quality and waste, their outputs are seldom linked to real-time MORL systems for automated environmental management. This remains a promising area for future exploration. Comparative evaluations suggest that RL-based demand response mechanisms reduce energy costs by up to 10–15%, yet no existing system actively links CV-based environmental monitoring (e.g., air quality, waste) to adaptive RL policies. This gap presents a promising direction for holistic smart environment solutions.

Future RL-environment models could leverage CV-based waste detection or air quality monitoring, transforming perceptual data into actionable RL-driven resource allocation strategies.

3.3.5 Smart living

Smart living is about improving public safety, healthcare systems, and emergency systems using intelligent vision-based systems for rapid urbanization. These systems leverage real-time data streams and predictive analytics to support faster incident detection, efficient public services, and citizen-centric city planning.

The Hybrid RL-Variable Neighborhood Strategy Adaptive Search (H-RL-VaNSAS) algorithm is utilized to develop a multi-objective model for creating sustainable, resilient, safe, and efficient urban bus routes. The model provides better-optimized bus routes with more resilience, sustainability, and safety to the cities of Ubon Ratchathani and Warinchamrab, given possible constraints in accessibility and the need to minimize travel distance. This approach contributes to urban transportation for smart living, where it provides a more systemic and realistic solution to some modern problems in relation to tourism areas [65]. By combining local search heuristics with deep RL policy optimization, the H-RL-VaNSAS algorithm adapts to spatial-temporal fluctuations in transport demand.

Similarly, multi-objective RL also has the following beneficial impacts on healthcare systems. A Generative Adversarial Network (GAN) is applied to the patient data, creates an artificial treatment scenario through Deep Neural Networks (DNN) and determines the transfer likelihood of individualized radiotherapy. This GAN-DNN integration enhances personalized treatment planning by generating synthetic clinical outcomes for reinforcement-based decision learning. An intelligent non-convex optimization problem of deep RL is employed to reduce environmental emissions in urban transport systems. The method also incorporates smart healthcare technologies to forecast the likelihood of appropriate travelling modes. The approach optimizes decisions and stimulates the commuter mode shift towards active transport, consequently enhancing health-related statistics and lowering gCO2 emissions [74], [76].

With the utility interconnection for integration of RL and storage system, the amount of greenhouse gas emissions and fossil diesel is minimized in microgrids in smart living. Propagation of a multi-microgrid power system can help to resolve security and privacy issues. An intelligent energy management approach is used depending upon preference-based multi-objective reinforcement learning (PMORL) methods for developing such systems [66]. PMORL frameworks enable energy optimization while incorporating user-defined utility preferences, offering an ethical and sustainable foundation for decentralized urban energy planning. In smart living, vision-based systems are commonly used for patient monitoring and public safety, yet RL agents operate in isolation from these perceptual systems. Integrating CV outputs into MORL-based adaptive healthcare and safety interventions is a critical research opportunity.

Table 6 presents an expanded synthesis of the reviewed smart city applications, highlighting both shared patterns and domain-specific distinctions in the integration of reinforcement learning (RL) and computer vision (CV). Across most domains, RL is predominantly used for policy optimization, while CV serves as a perception layer; however, deployment strategies differ based on task urgency, data modalities, and ethical constraints. For example, mobility systems demand real-time, low-latency control, whereas infrastructure domains prioritize long-term optimization of energy and resources. Smart governance raises concerns around privacy and surveillance, while healthcare and smart living applications emphasize human-centered and interpretable decision-making. This integrated taxonomy clarifies how multi-objective RL and CV are currently aligned—and where critical integration gaps persist. Notably, recent studies show that Pareto-optimal RL (PMORL) applied to healthcare logistics can reduce carbon emissions by up to 20% while improving patient activity recognition; however, CV in these cases often remains a standalone monitoring tool rather than being jointly optimized with RL. Addressing this disconnect through **vision-reinforced policy learning frameworks** represents a key opportunity for advancing next-generation smart living solutions

While RL is widely applied for optimizing energy usage and emissions, CV contributes to environmental monitoring by analyzing satellite imagery, thermal patterns, or pollution dispersion. The integration of CV-derived inputs into MORL pipelines enables more informed, adaptive control strategies in environmental management.

Across the reviewed domains, certain commonalities and divergences emerge. For instance, multi-agent RL is predominantly used in mobility for dynamic traffic control and lane optimization, while governance applications emphasize decentralized decision-making and policy simulation. CV tasks such as object detection and semantic segmentation are more prevalent in mobility and surveillance but are less integrated into infrastructure or energy-focused RL pipelines. While mobility prioritizes real-time adaptation and latency reduction, domains like infrastructure and environment favor long-term optimization and energy efficiency. Table 7 provides a synthesized taxonomy capturing these overlaps and distinctions. Patient activity recognition from CV systems can serve as input for RL agents managing healthcare logistics and emergency response, forming a cohesive perception-to-decision pipeline.

4 Discussion and research implications

This comprehensive review reveals a clear research trend toward applying reinforcement learning (RL), particularly multi-objective RL (MORL), and computer vision (CV) in various smart city domains. However, a detailed examination of the reviewed studies highlights that these technologies are often implemented in parallel rather than in integrated frameworks. While CV is primarily used for perception tasks such as surveillance, traffic monitoring, and environmental sensing, MORL focuses on optimizing decision-making processes, energy management, and resource allocation. A critical finding is the lack of unified systems that couple real-time visual data with adaptive RL-based control in urban environments. For example, in smart mobility, multi-camera tracking systems enhance surveillance, but traffic flow optimization typically operates independently of vision-based feedback. Similarly, in smart infrastructure, energy management relies on MORL algorithms, but visual occupancy data from CV systems is rarely incorporated into RL-driven HVAC or lighting control loops.

Table 6: Multi-objective RL and smart city applications

Studies	Multi-Objective Algorithm/Approach	Smart City Domain	Smart City Applications
[56], [61], [62], [63]	Multi-Agent Actor-Critic (MA2C) approach Multi-Objective Multiple Camera Tracking (MOMCT)	Smart Mobility	Traffic management, autonomous vehicles, public transport optimization, ITS, smart parking, traffic routing, and traffic police scheduling.
[41], [68], [70], [71]	Gaussian Naive Bayes K-Neighbors Feed-forward multilayer perceptron (MLP) neural network	Smart Infrastructure	Smart grids, smart building management, energy-efficient behavior and management optimization, renewable power-based energy storage and infrastructure monitoring

	Improved Whale Optimization Algorithm (IWOA)		
[74]	Swarm RL approach	Smart Governance	Digital identity systems, RFIDs, and intelligent surveillance systems.
[74], [75]	Q-learning RL-CNN technique Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNNs)	Smart Environment	Smart grids, Peer to Peer (P2P) energy system
[65], [66]	Hybrid RL-Variable Neighborhood Strategy Adaptive Search (H-RL-VaNSAS) Generative Adversarial Network (GAN) Deep Neural Networks (DNN) Preference-based multi-objective reinforcement learning (PMORL)	Smart Living	Sustainable, safe, and efficient urban bus routes, urban transport systems, smart healthcare and smart energy management systems.

Table 7: Taxonomy of RL and CV Applications in Smart City Domains.

Smart City Domain	Commonalities in RL/CV Use	Distinctive Aspects / Divergences
Smart Mobility	Multi-agent RL for dynamic control, CV for traffic monitoring and object tracking	Real-time decision-making, integration of sensor fusion, route optimization under traffic constraints
Smart Infrastructure	RL for energy/resource management, CV for environment monitoring	Focus on energy-saving, building-level automation, less real-time interaction compared to mobility
Smart Governance	RL for adaptive urban policy, CV for surveillance and digital twins	Emphasis on social systems, ethical implications, use of digital identity and civic engagement data
Smart Environment	RL for emission control, CV for anomaly detection in environmental data	Primarily focuses on predictive control, less integration with human-centric data
Smart Living	RL for healthcare logistics and emergency response, CV for health monitoring	Human-centered applications, ethical considerations in data privacy and health predictions

Moreover, the reviewed literature shows inconsistent treatment of "multi-objective" frameworks. Some studies refer to multi-criteria trade-offs like energy efficiency vs. comfort, while others focus on multi-agent cooperation without explicitly managing conflicting objectives. This indicates a need for clearer definitions and standardized methodologies when applying MORL in smart city contexts. Emerging technologies offer promising solutions to these gaps. The rise of digital twins enables simulation environments where CV and RL models can co-evolve, testing scenarios that combine perception and policy adaptation. Likewise, federated learning can support privacy-preserving CV data sharing across urban systems, while edge computing reduces latency, enabling real-time RL-CV interactions.

Despite advances, ethical considerations remain underexplored. Combining vision systems with autonomous decision-making raises concerns about bias, privacy, and surveillance overreach. Future research must address these societal implications alongside technical developments.

4.1 Research implications

- **Integration of RL and CV Pipelines:** Future studies should develop frameworks where CV directly informs RL reward functions or state representations, creating closed-loop adaptive systems.
- **Standardized Benchmarks:** The field lacks consistent performance metrics for evaluating RL-CV synergy in smart cities. Establishing such benchmarks would facilitate cross-domain comparisons.
- **Scalability and Deployment:** There is a need for real-world deployments at city-scale, leveraging edge computing and digital twin environments for testing RL-CV integration under dynamic conditions.
- **Ethical and Social Dimensions:** Addressing privacy, fairness, and transparency in vision-augmented decision-making is critical for public trust and adoption.

4.2 Emerging research directions and ethical considerations

The evolution of smart city technologies is opening new avenues for integrating reinforcement learning (RL) and computer vision (CV). Emerging trends include the use of Digital Twins for simulating urban environments, allowing RL and CV systems to co-adapt in virtual testbeds before real-world deployment. Additionally, edge computing is enabling real-time CV and RL decision-making at the sensor level, reducing latency and improving scalability [77]. The integration of Large Language Models (LLMs), such as GPT-4, with RL and CV pipelines presents further opportunities, particularly for dynamic policy generation and adaptive urban management. Moreover, federated learning is increasingly used to preserve data privacy while enabling cross-organizational learning across cities [78].

Alongside these advancements, ethical considerations must be addressed. The deployment of vision-based systems coupled with autonomous RL decision-making introduces risks of bias in surveillance systems, privacy violations, and lack of transparency in policy decisions. For example, facial recognition technologies deployed in urban surveillance have raised concerns about racial bias and wrongful identification, underscoring the need for bias mitigation protocols and public oversight. Responsible development requires frameworks for explainability, fairness, and data governance to ensure societal acceptance and long-term sustainability of these technologies [79].

5 Conclusion

The smart city concept is based on gathering real-time data from urban infrastructures like waste disposal and power grid systems. This data forms the foundation for adaptive feedback loops and real-time policy enforcement mechanisms in urban ecosystems. In the advancement of smart city infrastructure development, the role of multi-objective RL is essential and associated with more significant time and cost efficiency to analyze data-intensive workloads. The role of multi-objective RL and CV is substantial in a set of multi-objective applications in smart cities, such as cyber-security, healthcare, smart grids, smart waste management, smart traffic monitoring, security surveillance, and others. These techniques contribute to multi-criteria optimization under uncertainty, offering adaptive and scalable solutions that surpass conventional static models. Advancements in multi-objective digital twin technology have enabled governments to integrate smart city applications that monitor urban management among people, the environment, and infrastructure through digital identity systems, as well as civic engagement through intelligent surveillance. This convergence of RL, CV, and digital twin systems signals a shift toward predictive governance and citizen-driven service personalization in next-generation urban design.

6 Limitations and directions for future research

This comprehensive review explored that multi-objective RL in smart city applications are limited to few aspects such as data dependency, computational complexity, scalability issues, privacy and ethical concerns and trade-off management. The dependence on large-scale labeled data and real-time feedback loops poses significant barriers for deployment in data-scarce or privacy-restricted urban environments. However, future research can advance the analysis but exploring the high-quality datasets for training dealing with the unresolved challenges of infrastructure limitations using quantum computing, digital twins and Internet of Things. Emerging paradigms like federated learning and edge computing may also help mitigate data centralization and latency issues while preserving privacy. The focus of future researchers should depend upon combining multi-objective RL with adaptive ML models and optimization techniques to improve collaboration among multiple agents in smart cities. Emerging paradigms like federated learning and edge computing may also help mitigate data centralization and latency issues while preserving privacy.

References

[1] M. Herath and H. Mittal, “Adoption of artificial intelligence in smart cities: A comprehensive review,” *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100076, 2022. https://doi.org/10.1016/j.jjimei.2022.100076

[2] B. Song, F. Xing, H. Wang, X. Luo, S. Dai, P. Xiao, and Z. Zhao, “Evolutionary multi-objective reinforcement learning-based trajectory control and task offloading in UAV-assisted mobile edge computing,” *IEEE Transactions on Mobile Computing*, vol. 22, no. 12, pp. 7387–7405, 2022. https://doi.org/10.1109/tmc.2022.3208457

[3] H. Yang, J. Zhang, and J. Wang, “Urban traffic control in software-defined Internet of Things via a multi-agent deep reinforcement learning approach,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3742–3754, 2020. https://doi.org/10.1109/tits.2020.3023788

[4] D. Silver et al., “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, 2016. https://doi.org/10.1038/nature16961

[5] A. Dosovitskiy et al., “An image is worth 16×16 words: Transformers for image recognition at scale,” *arXiv preprint* arXiv:2010.11929, 2020.https://doi.org/10.48550/arXiv.2010.11929
Z. Zheng, Y. Zhou, Y. Sun, Z. Wang, B. Liu, and K. Li, “Applications of federated learning in smart cities: Recent advances, taxonomy, and open challenges,” *Connection Science*, vol. 34, no. 1, pp. 1–28, 2022. https://doi.org/10.1080/09540091.2021.1936455

[6] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006. https://doi.org/10.1191/1478088706qp063oa
Y. Zheng, Q. Hao, J. Wang, C. Gao, J. Chen, D. Jin, and Y. Li, “A survey of machine learning for urban decision making: Applications in planning, transportation, and healthcare,” *ACM Computing Surveys*, vol. 57, no. 4, art. 99, pp. 1–41, 2024. https://doi.org/10.1145/3695986

[7] M. J. Page et al., “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *BMJ*, vol. 372, p. n71, 2021.https://doi.org/10.1136/bmj.n71
D. Kothadiya, A. Chaudhari, R. Macwan, K. Patel, and C. Bhatt, “The convergence of deep learning and computer vision: Smart city applications and research challenges,” *Proceedings of the 3rd International Conference on Integrated Intelligent Computing, Communication and Security (ICIIC 2021)*, 2021. https://doi.org/10.2991/ahis.k.210913.003

[8] [11] U.-X. Lau, B. P. L. Marakkalage, S. H. Zhou, Y. Hassan, N. U. Yuen, C. Zhang, M. Tan, “A survey of data fusion in smart city applications,” *Information Fusion*, vol. 52, pp. 357–374, 2019. https://doi.org/10.1016/j.inffus.2019.05.004

[9] N. Buch, S. A. Velastin, and J. Orwell, “A review of computer vision techniques for the analysis of urban traffic,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 920–939, 2011. https://doi.org/10.1109/TITS.2011.2119372

[10] S. Mao, J. Sun, Q. Wang, X. Muthu, and B. Krishnamoorthy, “The importance of public support in the implementation of green transportation in the smart cities,” *Computational Intelligence*, vol. 40, no. 1, e12326, 2024. https://doi.org/10.1111/coin.12326

[11] G. Paiva, S. Ahad, M. A. Tripathi, G. Feroz, N. Casalino, “Enabling technologies for urban smart mobility: Recent trends, opportunities, and challenges,” *Sensors*, vol. 21, no. 6, p. 2143, 2021. https://doi.org/10.3390/s21062143

[12] A. Šurdonja, S. Giuffrè, and T. Deluka-Tibljaš, “Smart mobility solutions – necessary precondition for a well-functioning smart city,” *Transportation Research Procedia*, vol. 45, pp. 604–611, 2020. https://doi.org/10.1016/j.trpro.2020.03.051

[13] K. L. Ho, G. T. S. Tsang, Y. P. Wu, C. H. Wong, W. H. Choy, and W. H. Choy, “A computer vision-based roadside occupation surveillance system for intelligent transport in smart cities,” *Sensors*, vol. 19, no. 8, p. 1796, 2019. https://doi.org/10.3390/s19081796

[14] S. I. Musa, A. A. Malami, S. I. Alanazi, F. Ounaies, W. Alshammari, M. Haruna, “Sustainable traffic management for smart cities using Internet-of-Things-oriented intelligent transportation systems (ITS): Challenges and recommendations,” *Sustainability*, vol. 15, no. 13, p. 9859, 2023.https://doi.org/10.3390/su15139859

[15] Y. Zhang, T. Cheng, and J. Zou, “Multimodal transportation routing optimization based on multi-objective Q-learning under time uncertainty,” *Complex & Intelligent Systems*, vol. 10, pp. 3133–3152, 2024. https://doi.org/10.1007/s40747-023-01308-9

[16] R. K. Upadhyay, S. K. Sharma, and V. Kumar, “Introduction to intelligent transportation system and advanced technology,” in *Intelligent Transportation System and Advanced Technology*, R. K. Upadhyay, S. K. Sharma, and V. Kumar (eds), *Energy, Environment, and Sustainability*, Springer, Singapore, 2024. https://doi.org/10.1007/978-981-97-0515-3_1

[17] H. Akram, A. Abbas, S. Khan, M. Athar, A. Ghazal, and T. Al Hamadi, “Smart energy management system using machine learning,” *Computers, Materials & Continua*, vol. 78, no. 1, pp. 1–11, 2024. https://doi.org/10.32604/cmc.2023.032216

[18] S. Bahaghighat, M. Abedini, F. Xin, Q. Zanjireh, and M. M. Mirjalili, “Using machine learning and

computer vision to estimate the angular velocity of wind turbines in smart grids remotely," *Energy Reports*, vol. 7, pp. 8561–8576, 2021. https://doi.org/10.1016/j.egyr.2021.07.077

[19] H. C. Y. Chan and L. Chan, "Smart library and smart campus," *Journal of Service Science and Management*, vol. 11, no. 6, pp. 543–564, 2018.https://doi.org/10.4236/jssm.2018.116037
A. Dimitriadou and E. Lanitis, "A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms," *Smart Learning Environments*, vol. 10, no. 1, p. 12, 2023. https://doi.org/10.1186/s40561-023-00231-3

[20] S. Al-Fraihat, M. Joy, R. Masa'deh, and J. Sinclair, "Evaluating e-learning systems success: An empirical study," *Computers in Human Behavior*, vol. 102, pp. 67–86, 2020. https://doi.org/10.1016/j.chb.2019.08.004

[21] D. Mshali, H. Lemlouma, T. Moloney, and M. Magoni, "A survey on health monitoring systems for health smart homes," *International Journal of Industrial Ergonomics*, vol. 66, pp. 26–56, 2018. https://doi.org/10.1016/j.ergon.2018.02.002

[22] Z. Brommeyer, M. Whittaker, M. Mackay, M. Ng, and F. Liang, "Building health service management workforce capacity in the era of health informatics and digital health – A scoping review," *International Journal of Medical Informatics*, vol. 169, p. 104909, 2023. https://doi.org/10.1016/j.ijmedinf.2022.104909

[23] N. Jaya Lakshmi and N. Jabalia, "Body sensor networks as emerging trends of technology in health care systems: Challenges and future," in *Efficient Data Handling for Massive Internet of Medical Things Healthcare Data Analytics*, pp. 129–157, 2021. https://doi.org/10.1007/978-3-030-66633-0_6

[24] A. O. Okolo, C. A. Arowoogun, J. O. Chidi, and R. Adeniyi, "Telemedicine's role in transforming healthcare delivery in the pharmaceutical industry: A systematic review," *World Journal of Advanced Research and Reviews*, vol. 21, no. 2, pp. 1836–1856, 2024. https://doi.org/10.30574/wjarr.2024.21.2.0609

[25] T. Shaik, X. Tao, N. Higgins, L. Li, R. Gururajan, X. Zhou, and U. R. Acharya, "Remote patient monitoring using artificial intelligence: Current state, applications, and challenges," *WIREs Data Mining and Knowledge Discovery*, vol. 13, no. 2, e1485, 2023. https://doi.org/10.1002/widm.1485

[26] E. J. Peralta-Ochoa, A. M. Chaca-Asmal, P. A. Guerrero-Vásquez, L. F. Ordoñez-Ordoñez, J. O. Coronel-González, "Smart healthcare applications over 5G networks: A systematic review," *Applied Sciences*, vol. 13, no. 3, p. 1469, 2023. https://doi.org/10.3390/app13031469

[27] D. Z. Ribeiro, D. A. Melgarejo, D. C. Saadi, M. Rosa, R. L. Rodríguez, "A novel deep deterministic policy gradient model applied to intelligent transportation system security problems in 5G and 6G network scenarios," *Physical Communication*, vol. 56, p. 101938, 2023. https://doi.org/10.1016/j.phycom.2022.101938

[28] Q. Zhang, K. Xu, H. Pan, B. Zheng, "Explicit coordinated signal control using soft actor–critic for cycle length determination," *IET Intelligent Transport Systems*, 2024. https://doi.org/10.1049/itr2.12519

[29] L. D. Nguyen, K. K. Duong, T. Q. Vien, N. A. Le-Khac, and N. A. Nguyen, "Distributed deep deterministic policy gradient for power allocation control in D2D-based V2V communications," *IEEE Access*, vol. 7, pp. 164533–164543, 2019. https://doi.org/10.1109/access.2019.2952411

[30] G. Chen and X. Liu, "Federated deep reinforcement learning-based task offloading and resource allocation for smart cities in a mobile edge network," *Sensors*, vol. 22, no. 13, p. 4738, 2022. https://doi.org/10.3390/s22134738

[31] B. Qiu, C. Hu, Y. Chen, and Y. Zeng, "Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8577–8588, 2019. https://doi.org/10.1109/jiot.2019.2921159

[32] R. Duhayyim, M. A. Elfadil Eisa, T. A. Al-Wesabi, F. N. Abdelmaboud, A. Hamza, M. A. Zamani, A. S. Rizwanullah, and M. Marzouk, "Deep reinforcement learning enabled smart city recycling waste object classification," *Computers, Materials & Continua*, vol. 71, no. 3, 2022. https://doi.org/10.32604/cmc.2022.024431

[33] W. Ryu and K. Kim, "Multi-objective optimization of energy saving and throughput in heterogeneous networks using deep reinforcement learning," *Sensors*, vol. 21, no. 23, p. 7925, 2021. https://doi.org/10.3390/s21237925

[34] E. Vamvakas, D. Michailidis, P. Korkas, and C. Kosmatopoulos, "Review and evaluation of reinforcement learning frameworks on smart grid applications," *Energies*, vol. 16, no. 14, p. 5326, 2023. https://doi.org/10.3390/en16145326

[35] M. Vijayalakshmi and V. Saravanan, "Reinforcement learning-based multi-objective energy-efficient task scheduling in fog-cloud industrial IoT-based systems," *Soft Computing*, vol. 27, no. 23, pp. 17473–17491, 2023. https://doi.org/10.1007/s00500-023-09159-9

[36] J.-H. Yun, W. J. Park, S. Kim, J. Shin, M. Jung, S. Mohaisen, D. A. Kim, "Cooperative multi-agent deep reinforcement learning for reliable surveillance via autonomous multi-UAV control," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 7086–7096, 2022. https://doi.org/10.1109/tii.2022.3143175

[37] X. Li and W. Xu, "A hybrid evolutionary and machine learning approach for smart building: Sustainable building energy management design," *Sustainable Energy Technologies and Assessments*, vol. 65, p. 103709, 2024. https://doi.org/10.1016/j.seta.2024.103709

[38] S.-H. Quang Tran and D. Bae, "Proximal policy optimization through a deep reinforcement learning framework for multiple autonomous vehicles at a non-signalized intersection," *Applied Sciences*, vol.

10, no. 16, p. 5722, 2020. https://doi.org/10.3390/app10165722
[39] N. Mazzi, Y. Sassi, H. B. Errahimi, F. Es-Sbai, "A soft actor-critic reinforcement learning framework for optimal energy management in electric vehicles with hybrid storage," *Journal of Energy Storage*, vol. 99, p. 113344, 2024. https://doi.org/10.1016/j.est.2024.113344
[40] R. Hafiz, A. M. Parah, S. A. Bhat, "Reinforcement learning applied to machine vision: State of the art," *International Journal of Multimedia Information Retrieval*, vol. 10, no. 2, pp. 71–82, 2021. https://doi.org/10.1007/s13735-021-00209-2
[41] M. H. Abid, M. S. Apon, H. J. Hossain, S. Ahmed, A. Ahshan, R. Lipu, "A novel multi-objective optimization-based multi-agent deep reinforcement learning approach for microgrid resources planning," *Applied Energy*, vol. 353, p. 122029, 2024. https://doi.org/10.1016/j.apenergy.2023.122029
[42] N. Emami, L. Pacheco, A. Di Maio, and T. Braun, "RC-TL: Reinforcement convolutional transfer learning for large-scale trajectory prediction," in *NOMS 2022—2022 IEEE/IFIP Network Operations and Management Symposium*, 2022. https://doi.org/10.1109/noms54207.2022.9789883
[43] K. Diyan, M. Silva, and B. N. Han, "A multi-objective approach for optimal energy management in smart homes using reinforcement learning," *Sensors*, vol. 20, no. 12, p. 3450, 2020. https://doi.org/10.3390/s20123450
[44] C. E. Leite, G. M. C. Jiménez-Fernández, S. Salcedo-Sanz, C. G. Marcelino, and C. G. Pedreira, "Solving an energy resource management problem with a novel multi-objective evolutionary reinforcement learning method," *Knowledge-Based Systems*, vol. 280, p. 111027, 2023. https://doi.org/10.1016/j.knosys.2023.111027
[45] L. Prawiyogi, A. G. Purnama, S. Meria, "Smart cities using machine learning and intelligent applications," *International Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 102–116, 2022. https://doi.org/10.33050/italic.v1i1.204
[46] R. Ullah, Z. Al-Turjman, F. Mostarda, and L. Gagliardi, "Applications of artificial intelligence and machine learning in smart cities," *Computer Communications*, vol. 154, pp. 313–323, 2020. https://doi.org/10.1016/j.comcom.2020.02.069
[47] Z. Vázquez-Canteli, J. R. Ulyanin, S. Kämpf, and J. Nagy, "Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities," *Sustainable Cities and Society*, vol. 45, pp. 243–257, 2019. https://doi.org/10.1016/j.scs.2018.11.021
[48] V. Zhang, M. Wang, X. Sathishkumar, and V. Sivakumar, "Machine learning techniques based on security management in smart cities using robots," *Work*, vol. 68, no. 3, pp. 891–902, 2021.https://doi.org/10.3233/WOR-203423
[49] M. Shikunov and A. I. Panov, "Hierarchical reinforcement learning approach for the road intersection task," in *Biologically Inspired Cognitive Architectures 2019: Proceedings of the Tenth Annual Meeting of the BICA Society*, Springer, 2020. https://doi.org/10.1007/978-3-030-25719-4_64
[50] [54] O. A. Borges, D. F. Leite, J. P. R. Moreira, E. M. Carpinteiro, "Traffic light control using hierarchical reinforcement learning and options framework," *IEEE Access*, vol. 9, pp. 99155–99165, 2021. https://doi.org/10.1109/access.2021.3096666
[51] L. Mao, F. Li, Z. Lin, Y. Li, "Mastering arterial traffic signal control with multi-agent attention-based soft actor-critic model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, pp. 3129–3144, 2022. https://doi.org/10.1109/TITS.2022.3229477
[52] H. M. Louati, A. Louati, H. Kariri, E. Neifar, W. Hassan, M. K. Khairi, M. H. Farahat, M. A. El-Hoseny, "Sustainable smart cities through multi-agent reinforcement learning-based cooperative autonomous vehicles," *Sustainability*, vol. 16, no. 5, p. 1779, 2024. https://doi.org/10.3390/su16051779
[53] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint* arXiv:1509.02971, 2016. https://doi.org/10.48550/arXiv.1509.02971
[54] W. Jia and M. Ji, "Multi-agent deep reinforcement learning for large-scale traffic signal control with spatio-temporal attention mechanism," *Applied Sciences*, vol. 15, no. 15, p. 8605, 2025. https://doi.org/10.3390/app15158605
[55] H. Wei, G. Zheng, H. Yao, and Z. Li, "IntelliLight: A reinforcement learning approach for intelligent traffic light control," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ACM, 2017, pp. 2496–2505. https://doi.org/10.1145/3219819.3220096
[56] F. Çakıl and N. Aksoy, "Reinforcement learning-based multi-objective smart energy management for electric vehicle charging stations with priority scheduling," *Energy*, vol. 322, p. 135475, 2025. https://doi.org/10.1016/j.energy.2025.135475
[57] B. Fei and L. Han, "Multi-object multi-camera tracking based on deep learning for intelligent transportation: A review," *Sensors*, vol. 23, no. 8, p. 3852, 2023. https://doi.org/10.3390/s23083852
[58] E. Nisyrios and K. Gkiotsalitis, "Optimization-based approaches for traffic and fleet management of connected and autonomous vehicles: A systematic literature review," *International Journal of Intelligent Transportation Systems Research*, 2025. https://doi.org/10.1007/s13177-025-00536-2
[59] A. Syed, A. S. Sierra-Sosa, D. Kumar, and A. Elmaghraby, "Making cities smarter—optimization problems for IoT-enabled smart city development: A mapping of applications, objectives, constraints," *Sensors*, vol. 22, no. 12, p. 4380, 2022. https://doi.org/10.3390/s22124380
[60] M. T. Rafique, A. Mustafa, and H. Sajid, "Reinforcement learning for adaptive traffic signal

control: Turn-based and time-based approaches to reduce congestion," *arXiv preprint arXiv:2408.15751*, 2024. https://doi.org/10.48550/arXiv.2408.15751

[61] P. Sriprateep, K. Pitakaso, R. Khonjun, S. Srichok, T. Luesak, P. Gonwirat, S. Kaewta, C. Kosacka-Olejnik, and M. Enkvetchakul, "Multi-objective optimization of resilient, sustainable, and safe urban bus routes for tourism promotion using a hybrid reinforcement learning algorithm," *Mathematics*, vol. 12, no. 14, p. 2283, 2024. https://doi.org/10.3390/math12142283

[62] A. M. Kazmi, S. A. Khan, Z. Khan, A. Mazzara, and M. Khattak, "Leveraging deep reinforcement learning and healthcare devices for active travelling in smart cities," *IEEE Transactions on Consumer Electronics*, 2024. https://doi.org/10.1109/TCE.2024.3470978

[63] B. Nama, M. Nath, A. Bechra, N. Bhatia, J. Tanwar, S. Chaturvedi, and M. Sadoun, "Machine learning-based traffic scheduling techniques for intelligent transportation system: Opportunities and challenges," *International Journal of Communication Systems*, vol. 34, no. 9, e4814, 2021.https://doi.org/10.1002/dac.4814

[64] C. Konstantakopoulos, I. C. Barkan, A. R. He, S. Veeravalli, T. Liu, H. Liu, and C. Spanos, "A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure," *Applied Energy*, vol. 237, pp. 810–821, 2019. https://doi.org/10.1016/j.apenergy.2018.12.065

[65] H. Lin, Y. H. Lin, M. D. Tsai, K. T. Deng, and M. J. Ishii, "Multi-objective optimization design of green building envelopes and air conditioning systems for energy conservation and CO_2 emission reduction," *Sustainable Cities and Society*, vol. 64, p. 102555, 2021. https://doi.org/10.1016/j.scs.2020.102555

[66] M. A. Ghofrani, A. Nazemi, and S. D. Jafari, "HVAC load synchronization in smart building communities," *Sustainable Cities and Society*, vol. 51, p. 101741, 2019. https://doi.org/10.1016/j.scs.2019.101741

[67] S. Mason and K. Grijalva, "A review of reinforcement learning for autonomous building energy management," *Computers & Electrical Engineering*, vol. 78, pp. 300–312, 2019. https://doi.org/10.1016/j.compeleceng.2019.07.019

[68] W. L. Mocanu, E. Nguyen, P. H. Gibescu, and M. Kling, "Deep learning for estimating building energy consumption," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91–99, 2016. https://doi.org/10.1016/j.segan.2016.02.005

[69] Z.-J. M. Deng, T. Zhang, K. Shen, "A systematic review of a digital twin city: A new pattern of urban governance toward smart cities," *Journal of Management Science and Engineering*, vol. 6, no. 2, pp. 125–134, 2021. https://doi.org/10.1016/j.jmse.2021.03.003

[70] S. Tong, Z. Ye, F. Yan, M. Liu, and H. Basodi, "A survey on algorithms for intelligent computing and smart city applications," *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 155–172, 2021. https://doi.org/10.26599/bdma.2020.9020029

[71] K. Wolniak and R. Stecuła, "Artificial ntelligence in smart cities—applications, barriers, and future directions: A review," *Smart Cities*, vol. 7, no. 3, pp. 1346–1389, 2024. https://doi.org/10.3390/smartcities7030057

[72] M. Xu, J. Li, K. Abusara, "Preference-based multi-objective reinforcement learning for multi-microgrid system optimization problem in smart grid," *Memetic Computing*, vol. 14, no. 2, pp. 225–235, 2022. https://doi.org/10.1007/s12293-022-00357-w

[73] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017. https://doi.org/10.1109/MC.2017.9

[74] J. C. Jiang, B. Kantarci, S. Oktug, and T. Soyata, "Federated learning in smart city sensing: Challenges and opportunities," *Sensors*, vol. 20, no. 21, p. 6230, 2020. https://doi.org/10.3390/s20216230

[75] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, pp. 1–21, 2016. https://doi.org/10.1177/2053951716679679