International Journal of Research Publication and Reviews

A Critical Review of Edge AI for Real-Time Video Analytics in Smart Cities: Technological Foundations, Applications, Challenges, and Future Directions

Eshaan Warade[1], Prathmesh Sarode[2], Siddhant Yadav[3], Sayali Belhe[4]

[1]Department of Artificial Intelligence and Data Science, AISSMS Institute of Information Technology Pune, India eshaan.warade@gmail.com
[2]Department of Artificial Intelligence and Data Science AISSMS Institute of Information Technology Pune, India
prathmeshsarode18@gmail.com
[3]Department of Artificial Intelligence and Data Science AISSMS Institute of Information Technology Pune, India siddhantyadav@gmail.com
[4]Department of Artificial Intelligence and Data Science, AISSMS Institute of Information Technology Pune, India sayali.email@example.com

ABSTRACT

Abstract—Smart cities are increasingly adopting advanced technologies to enhance urban living, with real-time video analyt ics playing a pivotal role in applications such as traffic manage ment, public safety, and environmental monitoring. Traditional cloud-based video analytics face challenges in latency, bandwidth, and privacy, which Edge AI addresses by processing data locally at the network's edge. This review critically examines Edge AI's role in real-time video analytics for smart cities, exploring its technological foundations, applications, challenges, countermea sures, and future research directions. Drawing on recent studies and real-world implementations, the paper highlights how Edge AI is transforming urban management while emphasizing the need for responsible development to address ethical and technical concerns. Edge AI enables data processing closer to the data source, sig nificantly reducing latency and bandwidth consumption—critical factors in real-time video analytics. This localized processing ensures that data privacy is maintained, as sensitive information need not be transmitted over networks to centralized data centers. In the context of smart cities, this capability is invaluable for applications requiring immediate responses, such as traffic incident detection, crowd management, and emergency response coordination.

The technological underpinnings of Edge AI involve the integration of advanced machine learning algorithms with edge computing infrastructure. This synergy allows for the deploy ment of intelligent analytics directly on devices like surveillance cameras and traffic sensors. Recent advancements in hard ware accelerators, including Field-Programmable Gate Arrays (FPGAs), Application-Specific Integrated Circuits (ASICs), and Neural Processing Units (NPUs), provide the computational efficiency needed to run complex deep learning models within the power constraints of edge devices. These hardware innovations, coupled with model optimization techniques like quantization, pruning, and knowledge distillation, have dramatically reduced the resource requirements for deploying sophisticated computer vision algorithms at the edge. Intelligent urban edge AI applications are diverse and effective.

In traffic management, Edge AI systems can analyze live video feeds to detect congestion, accidents, or violations, allowing for dynamic traffic signal adjustments and rapid incident response. Cities likeCities like Singapore and Barcelona implement intel ligent traffic management systems that use edge-based video analytics to optimize traffic flow, reducing average pendulum time by 15-20 % and reducing carbon emissions through reduced idling. For public safety, real-time video analytics can identify unusual behaviors or unauthorized access in sensitive areas, en hancing surveillance capabilities without compromising privacy. Edge AI-powered systems in cities like Chicago and London can detect potential security threats such as abandoned objects, perimeter breaches, or crowd formations that might indicate unrest, triggering immediate responses from law enforcement or emergency services. Environmental monitoring also benefits, as Edge AI can process visual data to detect pollution levels, illegal dumping, or changes in green cover, contributing to urban sustainability efforts. In Seoul, edge-based cameras monitor air quality by analyzing visual indicators of pollution, while cities in the Netherlands use similar technology to detect illegal waste disposal in public spaces. These applications demonstrate how Edge AI ex tends beyond traditional security applications to address broader aspects of urban management and environmental stewardship. The healthcare dimension of smart cities has gained particular relevance in the wake of global health crises. Edge AI systems deployed in public spaces can monitor compliance with health protocols, detect potential disease, vectors, and facilitate contact tracing while preserving individual privacy. During the COVID 19 pandemic, several cities utilized edge-based video analytics to monitor mask usage and social distancing in public areas, providing valuable data for public health initiatives without compromising personal identification information. Despite its advantages, the implementation of Edge AI in smart cities presents several challenges.

Resource constraints on edge devices can limit the complexity of deployable algorithms, necessitating the development of lightweight models. Ensuring data security and privacy at the edge is paramount, requiring robust encryption and access control mechanisms. Additionally, the heterageneity of devices and the need for interoperability pose significant integration challenges. The ethical dimensions of widespread video analytics in urban spaces cannot be overlooked, with questions regarding surveillance, consent, and algorithmic bias requiring careful consideration as Edge AI technologies proliferate. To address these challenges,

ongoing research focuses on opti mizing machine learning models for edge deployment, developing standardized protocols for device interoperability, and enhancing security frameworks tailored for edge environments. Researchers are increasingly focusing on developing explainable AI models that provide transparency in decision-making processes, along with fairness-aware algorithms that minimize biases in video analysis. Future research directions include the exploration of federated learning approaches to enable collaborative model training without data sharing, and the development of adaptive systems capable of real-time learning and decision-making in dynamic urban environments.

In conclusion, Edge AI stands as a transformative force in the evolution of smart cities, offering solutions that align with the demands for efficiency, responsiveness, and privacy. Its role in real-time video analytics is pivotal, providing the means to process vast amounts of visual data swiftly and securely. The integration of Edge AI with emerging technologies such as 5G networks, Internet of Things (IoT) ecosystems, and digital twins promises to further enhance the capabilities of smart city infrastructure. As cities continue to grow and face complex challenges, the responsible development and deployment of Edge AI technologies will be crucial in shaping sustainable and resilient urban futures.

Introduction

Smart cities integrate advanced technologies to improve Residents' quality of life, maximize resource utilization, and., and promote sustainability [1]. Real-time video analytics, pow ered by widespread surveillance cameras, is central to these efforts, enabling applications like traffic flow optimization, public safety enhancement, and environmental monitoring. However, cloud-based video analytics often suffer from high latency, bandwidth constraints, and privacy risks due to cen tralized data processing [3]. Edge AI, which combines artificial intelligence with edge computing, processes data locally on devices like cameras or edge servers, mitigating these issues. By reducing latency, minimizing bandwidth usage, and enhancing privacy through localized processing, Edge AI is well-suited for smart city applications [1]. As urban populations grow and cities face increasing complexity, Edge AI's role in real-time video analytics becomes critical. This review provides a comprehensive analysis of Edge AI for real-time video analytics in smart cities, covering:

• Technological foundations enabling Edge AI systems

• Current applications across urban domains

• Technical, ethical, and operational challenges

• Countermeasures and safeguards to address limitations

• Next steps for expanding the knowledge base.

By integrating knowledge from scholarly studies and prac tical experience.-world deployments, this paper aims to guide researchers, urban planners, and policymakers in leveraging Edge AI responsibly to build smarter, more sustainable cities.

TECHNOLOGICAL FOUNDATIONS

Edge AI integrates artificial intelligence with edge com puting, enabling data processing on devices at the network's edge, such as cameras, sensors, or local servers [1]. This ap proach offers reduced latency, lower bandwidth requirements, enhanced privacy, and operability in resource-constrained en vironments, making it ideal for real-time video analytics in smart cities.

A.   Hardware Platforms

Edge AI relies on specialized hardware to execute complex AI models efficiently:

Google Edge TPU: A tensor processing unit optimized for machine learning inference on edge devices. • NVIDIAJetson Series: High-performance platforms like Jetson Nano and TX2, designed for AI-driven video analytics [5].• Intel Movidius VPU:A Vision Processing Unit Tailored. • Apple Neural Engine: Integrated into Apple devices for on-device AI processing. These platforms empower edge devices to manage computation, tasks like object detection and facial recognition in real time.

B.   Software Frameworks

Several software frameworks support the development and deployment of Edge AI applications: • NVIDIA Metropolis: A platform for building intelligent video analytics applications, leveraging IoT sensors [6]. • DeepStream SDK: Tools for creating AI-based video analytics pipelines. • TAO Toolkit: Facilitates training and deploying opti mized AI models on edge devices. • TensorRT: An SDK for high-performance deep learning inference [7]. These frameworks streamline the development of scalable, cloud-native solutions for smart city applications.

C.   Distributed Architectures

Advanced architectures like MELINDA enhance Edge AI's capabilities by distributing video analytics tasks across edge nodes [2]. MELINDA employs a three-tier architecture, divid ing processing into: • Measurement Level Tasks (MLT): Initial data collection and preprocessing. • Feature Level Tasks (FLT): Feature extraction using neural networks. • Decision Level Tasks (DLT): Final decision-making based on processed data. By using hardware

accelerators and load balancing, MELINDA achieves significant improvements in processing throughput, demonstrating the potential of distributed Edge AI systems

APPLICATIONS IN SMART CITIES

Edge AI for real-time video analytics supports a wide range of applications in smart cities, addressing urban challenges and enhancing operational efficiency [1].

A. Security and Surveillance

Edge AI analyzes video feeds from public spaces to detect anomalies, recognize faces, and identify suspicious behavior. For instance, the MELINDA system enables facial recogni tion in smart buildings, enhancing security while maintaining privacy through localized processing [2]. Such systems are deployed in airports, train stations, and shopping centers to ensure public safety.

B. Transportation and Traffic Management

Real-time traffic monitoring is a critical application, en abling dynamic traffic signal adjustments, accident detection, and smart parking solutions. A pilot project in Liverpool, Australia, deployed smart visual sensors using edge computing and deep neural networks to track multi-modal transportation while prioritizing citizen privacy [4]. These systems help reduce congestion and improve urban mobility.

C. Healthcare

In smart hospitals or public health scenarios, Edge AI monitors patient flows, ensures compliance with safety pro tocols, and detects incidents like falls in elderly care facilities, enhancing healthcare delivery in urban settings. The intro duction of edge AI into healthcare infrastructure introduces a new. Revolutionizing Medical Services Through Real-Time Data Pro. data processing, reducing latency, and improving the responsiveness of critical systems. As cities grow more densely populated, the demand for efficient, timely, and scal able healthcare solutions increases. Edge AI offers a promising approach by processing data close to the source—whether in a hospital room, ambulance, or care home—minimizing the delay associated with cloud computing. One prominent use case of Edge AI in smart healthcare is patient monitoring through wearable devices and smart sen sors.

These edge-enabled devices can continuously track Vital signs like heart rate, blood pressure, and oxygen saturation levels are important indicators of a person's health. By an alyzing data on-device or at nearby edge nodes, anomalies such as arrhythmias or respiratory distress, can be detected instantly, triggering alerts to medical staff. This immediacy can be lifesaving, especially in critical care units, postoperative recovery wards, or for patients with chronic illnesses. In elderly care, fall detection is another critical application. Edge AI-enabled vision systems installed in rooms or common areas can identify abnormal movements or sudden changes in posture that suggest a fall. Unlike traditional surveillance systems that require constant human monitoring or rely on cloud processing, Edge AI allows for faster decision-making and privacy preservation by keeping sensitive video data local to the device. Moreover, these systems can distinguish be tween false alarms—such as sitting down quickly—and actual emergencies, reducing unnecessary interventions.

Hospital workflow optimization is another area where Edge AI plays a significant role. By tracking staff movement and equipment usage through computer vision or RFID-based systems, hospitals can ensure better allocation of resources. For instance, locating critical equipment like defibrillators or ventilators in emergencies becomes faster and more efficient. Similarly, monitoring patient movement helps avoid conges tion in hallways, reduces waiting times, and enhances patient satisfaction. Furthermore, Edge AI supports infection control and com pliance monitoring, which became especially crucial during the COVID-19 pandemic. Smart cameras powered by edge computing can detect face mask usage, monitor social distanc ing, and count room occupancy in real-time. These systems ensure that safety protocols are adhered to without burdening human staff.

In operating rooms and intensive care units, Edge AI can monitor hand hygiene practices, thereby reducing the risk of hospital-acquired infections. Another emerging application is in telemedicine and remote diagnostics. In rural or underserved areas where high-speed Internet connection speed may be unstable, edge AI enables diagnostic devices to perform preliminary analysis before transmitting data to specialists. For example, edge-enabled ultrasound machines can assist field technicians by offering AI-based suggestions and pre-diagnosis support. This reduces the dependency on continuous connectivity and accelerates the delivery of care. Emergency response systems also benefit significantly from Edge AI.

Ambulances equipped with edge devices can begin patient assessment en route to the hospital, transmitting pro cessed summaries to emergency departments. This pre-arrival information helps doctors prepare for immediate intervention upon arrival, saving valuable minutes in trauma or cardiac cases. The integration of Edge AI into electronic health records (EHR) also enhances patient safety. By locally analyzing data from multiple systems— lab results, imaging, prescrip tions—edge nodes can flag potential issues such as allergic drug reactions or conflicting medications before the informa tion is even uploaded to centralized systems. This proactive approach contributes to safer and more effective treatment. Moreover, Edge AI promotes data privacy and security, which are paramount in healthcare. By processing and storing data locally, the risk of cyberattacks and data breaches asso ciated with cloud transmission is reduced. This is especially critical in regions with strict data protection regulations such as GDPR or HIPAA. In conclusion, Edge AI is transforming the healthcare land scape in smart cities by enabling real-time, context-aware, and secure healthcare solutions. From monitoring patients and improving elderly care to optimizing hospital operations and enhancing emergency services, the role of Edge AI is multifaceted and growing. As technology continues to evolve, the synergy between healthcare and edge computing will further enhance the quality, accessibility, and efficiency of medical services across urban environments.

D.   Education

Schools and universities use video analytics for safety mon itoring, attendance tracking, and behavior analysis, creating secure learning environments.

E.   Retail In commercial areas, Edge AI analyzes customer traffic to optimize store layouts, enhance security, and provide person alized shopping experiences, boosting economic activity in smart cities.

F.   Environmental and Energy Management Edge AI keeps track of environmental factors, like air quality or waste management, and adjusts energy consump tion in public buildings and street lighting, helping achieve sustainability objectives.

IV. CHALLENGES AND LIMITATIONS

Despite its potential, Edge AI for real-time video analytics faces several challenges that could hinder its adoption [1], [3].

A. Algorithm Accuracy and Robustness AI models must perform reliably across diverse urban environments, including varying lighting, weather, and camera qualities. Misclassifications or failures in complex scenarios can undermine system effectiveness. B. Power Consumption Edge devices have limited power resources, requiring energy-efficient algorithms and hardware to sustain continuous operation without frequent recharging or replacement.

C. Integration with Existing Infrastructure Retrofitting existing city infrastructure, such as legacy CCTV systems, with Edge AI technologies is complex and costly, requiring standardized protocols and careful planning.

D. Privacy and Security Processing sensitive video data locally increases the risk of data breaches or misuse if edge devices are not adequately secured. Privacy concerns are particularly significant in public spaces where surveillance is widespread.

E. Scalability 5 The proliferation of cameras in smart cities generates mas sive data volumes, challenging the scalability of edge-based systems. Managing these data streams without overwhelming edge nodes is a critical issue [3].

V. COUNTERMEASURES AND SAFEGUARDS

To address the challenges, several countermeasures and safeguards can be implemented:

A. Enhancing Algorithm Accuracy Advanced training techniques, such as transfer learning and ensemble methods, combined with diverse datasets, can improve model robustness across urban environments.

B. Reducing Power Consumption Model optimization techniques like quantization and prun ing reduce computational demands, enabling efficient opera tion on power-constrained edge devices [1].

C. Facilitating Integration Standardizing protocols and APIs ensures compatibility with existing infrastructure, simplifying the deployment of Edge AI systems in smart cities.

D. Protecting Privacy Techniques like data anonymization, federated learning, and secure multi-party computation keep sensitive data local, minimizing privacy risks [2].

E. Enhancing Security Regular software updates, data encryption, and robust au thentication mechanisms protect edge devices from cyberat tacks, ensuring system integrity. These countermeasures are critical for the responsible and effective deployment of Edge AI in urban settings.

VI. FUTURE RESEARCH DIRECTIONS

The field of Edge AI for real-time video analytics is evolving rapidly, with several research directions poised to enhance its capabilities

A. Interdisciplinary Collaboration Combining expertise from AI, computer vision, networking, and urban planning can develop holistic solutions tailored to smart city needs [1].

B. Advancements in Hardware and Software More powerful, energy-efficient edge devices and sophis ticated AI models will expand the scope of real-time video analytics, enabling more complex applications.

C. Integration with 5G The rollout of 5G networks offers high-speed, low-latency connectivity, enhancing Edge AI's performance in data intensive applications [8].

D. Decentralized AI Blockchain-based approaches for secure, transparent data sharing among edge devices could improve trust and interop erability in smart city systems.

E. Ethical and Regulatory Frameworks Developing guidelines for ethical AI use in public spaces is essential to address privacy concerns and gain public trust, particularly in surveillance-heavy applications.

F. Global Scalability Scaling successful pilot projects, such as the Liverpool traffic monitoring system, to larger cities and diverse regions will validate and refine Edge AI technologies [4].

VII. CONCLUSION

Real-time video analytics is revolutionizing smart cities by providing faster, more private, and efficient urban management capabilities. Its applications in security, transportation, health care, and environmental monitoring demonstrate its potential to address complex urban challenges. However, issues like algorithm accuracy, power consumption, privacy, and scal ability require ongoing attention. Through countermeasures like model optimization, standardized integration, and robust security, these challenges can be mitigated. Future research in hardware advancements, 5G integration, and ethical frame works will further enhance Edge AI's impact. By balancing technological innovation with responsible deployment, Edge AI can help create smarter, more sustainable cities for the future.

REFERENCES

[1] E. Badidi, K. Moumane, and F. El Ghazi, "Opportunities, Applications, and Challenges of Edge-AI Enabled Video Analytics in Smart Cities: A Systematic Review," IEEE Access, vol. 12, pp. 1–15, 2024. [Online]. Available: https://www.researchgate.net/publication/ 372835650 Opportunities Applications and Challenges of Edge-AI Enabled Video Analytics in Smart Cities A Systematic Review

[2] A. L. Rocha et al., "Leveraging Edge Intelligence for Video Analytics in Smart City Applications," Information, vol. 12, no. 1, 2021, [Online]. Available: https://www.mdpi.com/2078-2489/12/1/14.

[3] D. Li et al., "Real-Time Video Analytics: The Killer App for Edge Computing," Computer, vol. 50, no. 10, pp. 58–67, 2017, [Online]. Available: https://www.computer.org/publications/tech news/research/real-time-video-analytics-for-camera-surveillance-in edge-computing.

[4] "Edge-Computing Video Analytics for Real-Time Traffic Monitoring in a Smart City," Procedia Computer Science, 2024, [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC6540244/.

[5] "Edge AI For Smart Cities & Video Analytics," Arrow.com, 2021, [Online]. Available: https://www.arrow.com/en/research-and events/articles/edge-ai-and-video-analytics-for-smart-cities.

[6] "Metropolis: Intelligent Vision AI for Smart Infrastructure," NVIDIA, 2021, [Online]. Available: https://www.nvidia.com/en-us/autonomous machines/intelligent-video-analytics-platform/.

[7] "Building Smart Cities With Help of AI Video Analytics," NVIDIA, 2021, [Online]. Available: https://www.nvidia.com/en us/industries/smart-cities/.
[8] "Reviving 5G," Smart Smart Cities Cities with World, Edge 2021, Computing [Online]. and Available: https://www.smartcitiesworld.net/opinions/opinions/reviving-smart cities-with-edge-computing-and-5g