TOPICAL REVIEW

Opportunities, Applications, and Challenges of Edge-AI Enabled Video Analytics in Smart Cities: A Systematic Review

ELARBI BADIDI[1], KARIMA MOUMANE[2], AND FIRDAOUS EL GHAZI[3]

[1]Department of Computer Science and Software Engineering, CIT, UAE University, Al Ain, United Arab Emirates
[2]Software Project Management Research Team, ENSIAS, Mohammed V University, Rabat 8007, Morocco
[3]Physics and Applications Laboratory, Ibn Tofail University, Kenitra 14000, Morocco

Corresponding author: Elarbi Badidi (ebadidi@uaeu.ac.ae)

ABSTRACT Video analytics with deep learning techniques has generated immense interest in academia and industry, captivating minds with its transformative potential. Deep learning techniques and the deluge of video data enable the mechanization of tasks that were once the exclusive domain of human effort. Furthermore, edge intelligence is emerging as an interdisciplinary technology that drives the fusion of edge computing and artificial intelligence (AI). Edge computing allows the Internet of Things (IoT) devices with limited resources to offload their compute-intensive AI applications to the network edge servers for execution. Specifically, AI workloads for video analytics can be moved to the network edge from the cloud, providing improved latency and bandwidth savings, among other benefits. This article reviews current technologies used in Edge AI-assisted video analytics in smart cities. It examines the various artificial intelligence models and privacy-preserving techniques used in edge video analytics. It identifies the various applications of video analytics in smart cities, including security and surveillance, transportation and traffic management, healthcare, education, sports and entertainment, and many more. Besides, it highlights the challenges of edge video analysis and open research issues. It is expected that this review will be valuable for researchers, engineers, and decision-makers who want to understand the landscape and scale of edge video analytics in smart cities.

INDEX TERMS Artificial intelligence, deep learning, edge computing, edge intelligence, edge video analytics, machine learning, smart city.

I. INTRODUCTION

In recent years, *video analytics*, also known as video content analysis, has attracted growing interest in academia and industry. As a result of the vast quantities of video data that are easily accessible and the noteworthy improvements in deep learning methodologies, video analytics has enabled the automation of activities that were formerly exclusively handled by humans. Recent improvements in video analytics have changed the game. Video analytics applications range from real-time applications that monitor buildings, airports,

The associate editor coordinating the review of this manuscript and approving it for publication was Valentina E. Balas .

train stations, schools, and universities, as well as traffic congestion in cities and on highways, to detect specific events, such as car accidents or crowd stampedes, and trigger appropriate alerts. Other applications analyze customer traffic in retail stores to maximize sales, and other, more familiar scenarios include facial recognition and smart parking.

Deep Neural Networks (DNN) growing usage has made it possible to train video analytics systems that mimic human behavior, leading to a paradigm shift. It started with systems based on classical computer vision techniques that trigger an alarm when, for example, the camera image becomes too dark or changes drastically. It moved to systems that can identify specific objects in an image and track their path.

Video analytics software can run centrally on servers, usually in the monitoring station, or on cloud servers to take advantage of their processing power and unlimited storage capabilities. Alternatively, it can be embedded in the cameras or local servers, a strategy known as *edge processing*.

Edge computing is an emerging paradigm that moves traditional cloud-based network computing capabilities from centralized data centers to end-user devices and local area networks. It fundamentally extends the cloud computing architecture to the network edge, enabling an innovative variety of silent services and applications for end users. Since the advent of the implementation of IoT applications in diverse sectors, the quantity of Internet-enabled devices has surged from millions to billions. It is anticipated that this trend will persist in the near future.

Traditional centralized, cloud-based systems cannot respond to all connected devices in real time without compromising the user experience. To address the challenges of IoT infrastructure, edge computing has been introduced, promising low latency, high mobility, and broad geographic coverage with compensation for many nodes. Edge computing is still in the development phase. The surging prevalence of this technological advancement can be ascribed to its extensive capacity for a diverse array of applications, encompassing IoT-based frameworks, real-time computing systems, energy-efficient computing applications, latency-sensitive applications, and mobile applications. These applications have been studied extensively in various scholarly works such as [1], [2], [3], [4], and [5].

Smart cities promise more convenience and more services for citizens. Smart city projects today are essentially based on IoT infrastructures and edge devices. One critical problem they can solve is public safety [6]. For cities around the world, public safety is a growing concern Therefore, cities of the future should be equipped with edge computing technology to provide crime-fighting capabilities to police and emergency services and make cities safer [7], [8]. Safeguarding urban areas is extremely important for cities' growth and prosperity and citizens' well-being. Cities increasingly deploy security cameras to preserve public order in sensitive places such as airports, train stations, shopping centers, street intersections, and public spaces. AI is also helping to maintain security by being used to quickly identify trespassing or other such incidents based on analysis of large volumes of surveillance camera footage.

The field of edge AI-enabled video analytics is currently in its nascent stage, characterized by ongoing research endeavors focused on intelligent video analytics. This technology holds immense promise for catalyzing transformative changes across various sectors, encompassing but not restricted to public safety, smart buildings, healthcare, and transportation services. While extant literature, as evidenced by prior reviews [9], [10], [11], provides a comprehensive overview of research efforts in video analytics, a notable research gap persists pertaining to the extensive corpus of studies investigating AI-assisted video analytics. This review aims to fill this gap by thoroughly assessing the current scholarly works on edge video analytics, identifying the inherent challenges within this domain, and elucidating the potential benefits applications can derive from leveraging AI-assisted video analytics at the network edge.

A. CONTRIBUTIONS

The contribution of this review paper on edge video analytics can be summarized as follows:

1) Review of the current state of the art in edge video analytics, including key technologies, applications, and challenges.
2) Analysis of the benefits of edge video analytics over traditional cloud-based video analytics, including real-time processing, reduced latency, improved privacy and security, and ability to deploy applications in resource-limited environments.
3) Synthesis of the current research and development trends in edge video analytics, including recent advances in algorithms, hardware, and software solutions.
4) Identification of the key challenges that need to be addressed to realize the potential of edge video analytics fully. These challenges essentially concern improving the accuracy and robustness of the algorithms, reducing power consumption, and facilitating integration into existing video analytics systems.
5) Provision of insights and recommendations for future research and development in edge video analytics, highlighting the need for interdisciplinary collaboration and the importance of addressing the current challenges and limitations.

Through an in-depth analysis of the existing scholarly literature, this manuscript endeavors to enhance the current comprehension of the nascent area of edge video analytics and establish a foundation for forthcoming research and development.

B. STRUCTURE OF THE REVIEW

The remainder of this review is organized as follows: Section II presents video analytics and edge AI fundamentals and describes their potential to support smart city operations. Section III discusses the benefits of AI-assisted video analytics in general and edge video analytics in particular and highlights the various Edge AI platforms for video analytics. The methodology used in this review is described in Section IV. Section V describes the results of our investigation and details the taxonomy of the eligible works of this study. Section VI describes the most important applications and use cases of edge video analytics in smart cities. The challenges of edge video analytics and the open research issues are highlighted in Section VII. Finally, Section VIII concludes this review paper.

II. VIDEO ANALYTICS AND EDGE AI

A. VIDEO ANALYTICS OVERVIEW

Video analytics (VA) has revolutionized various domains, including traffic management, security operations, healthcare, and retail. By intelligently analyzing video data and CCTV footage, VA unveils hidden patterns and correlations, empowering informed decision-making and anticipating future events. It surpasses human monitoring in precision and efficiency, triggering timely alarms. VA also captures valuable business data, making it a vital resource for security managers and commercial sectors.

The escalating imperative for security measures has triggered a surge in the deployment of surveillance cameras across an expansive array of locations. Airports, train stations, highways, stadiums, public gatherings, schools, and supermarkets stand as examples where the prevalence of these watchful electronic sentinels has witnessed a substantial increase. Intelligent video surveillance aims to learn the usual events and detect uncommon events in the observed area. These unusual events differ from regular events and are called unusual or suspicious events [12], [13], [14]. They are unseen activities that occur infrequently and are not repeated.

To differentiate routine activities from potential threats, software operators create custom rulesets tailored to site-specific risk factors and specific incident types. These rules encompass factors like crowd behavior, loitering, and unusual movement to aid in event detection. Multiple rule variations are available, dependent on the chosen model, and ongoing development leads to the creation of new rules. This adaptability enables security managers to configure the software for targeted monitoring of specific areas or scenarios.

Video analysis utilizes a mathematical model of the background scene to detect objects based on pixel differences. This enables the detection of security breaches, high-risk zone entry, license plates, and object placement/removal. Alerts are promptly generated, either as video pop-ups, emails, or smartphone notifications, ensuring immediate action by security personnel. The key benefits of video analytics are:

- *Automated surveillance and increased efficiency.* Video analytics offers remarkable efficacy and precise outcomes, reducing the need for constant manual CCTV monitoring and extensive security personnel. Unlike humans, it operates tirelessly, detecting incidents that may go unnoticed. This invaluable solution allows organizations to save costs and time while providing continuous surveillance.
- *Increased likelihood of incident detection and prevention.* Video analytics empowers operators to swiftly respond to incidents using CCTV footage, potentially preventing crimes through early warnings. Real-time alerts enable security managers to make informed decisions by accessing live footage, facilitating dynamic responses like alarm triggering and immediate intervention. Additionally, video analytics aids in post-incident investigations by offering efficient query-based searches, saving time by quickly locating specific events, suspects, or relevant information without manual effort.
- *Cost-effective solution.* Accurate video analysis enables end users to potentially reduce the number of security guards, resulting in long-term cost savings. Additionally, the technology allows operators to selectively store footage of suspicious events, eliminating the need to store extensive hours of irrelevant footage and further reducing costs associated with storage.
- *Providing business intelligence dashboards.* Video analytics empowers organizations with valuable business intelligence through reports, dashboards, and heat maps. It reveals insights into daily activities, trends, and patterns across departments, including people counting, customer behavior, traffic tracking, and queue monitoring. Automation streamlines dashboard creation, focusing on active hours rather than inactive periods for optimized tracking.

Figure 1 illustrates the architecture of a typical video analytics engine, which includes the following components:

- *Video Inputs.* A video input is a system that captures and digitizes video data streams from IP cameras or video encoders, empowering the realm of video analytics.
- *Video Management System.* A system that manages the video inputs and makes the video streams available for analytics.
- *Object Detection and Tracking.* Algorithms that revolutionize video analytics by precisely identifying and tracing objects within dynamic video streams.
- *Event Detection.* Algorithms that analyze video streams to detect specific events such as motion and recognize faces and changes in scenes.
- *Video Indexing.* Indexing the video content based on metadata such as time-stamp, location, and object type.
- *Data Storage.* A data storage system to store the video content and metadata.
- *User Interface.* A user interface that displays the results of analysis of the video content. It also permits users to interact with the data for more insights.
- *Data Management.* A system like a database or data management system, which allows users to manage and retrieve the stored video data and analysis results.
- *Reporting and Analytics.* A system for generating reports and rendering the data in various ways.

B. EDGE AI OVERVIEW

Edge computing, sometimes called IoT, continues to spread and has become an essential part of most enterprise strategies in recent years [15], [16], [17]. IoT devices, sensors, and smartphones are transforming many businesses from the ground up. Nevertheless, the emergence of AI has a phenomenal impact on what happens at the edge. Increasing computing power at the edge, combined with the ease of using machine learning and deep learning, makes edge devices extremely
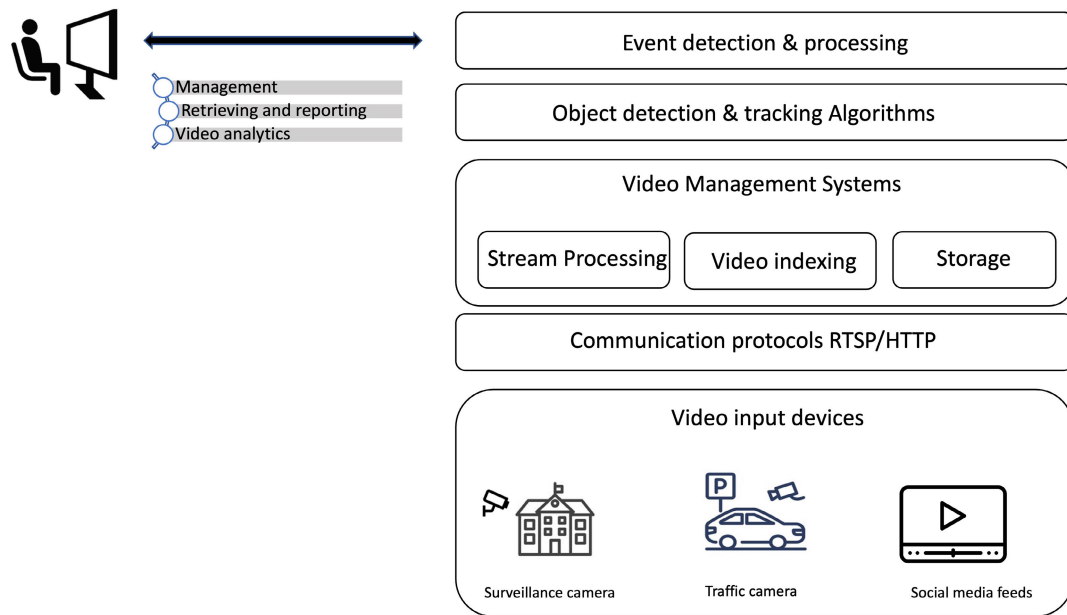
FIGURE 1. Architecture of a video analytics system.

intelligent [18], [19], [20], [21], [22]. This allows the devices to provide real-time insights and predictive analytics without sending data to remote cloud servers. Many intelligent business solutions are already being deployed in manufacturing. Modern factories use various industrial IoT devices to be alerted to problems in their supply chain and proactively avoid unplanned downtime [23], [24]. In smart cities, small devices on a roadside radar can now instantly detect a car speeding, the occupants in the car, and whether or not the driver has a license [25], [26], [27], [28].

AI with pre-trained models can empower smart city decision-makers by enabling them to make informed decisions that benefit the city and its citizens [29], [30]. For example, many smart city areas will benefit from two typical image processing-based tasks, image classification and object recognition, which occur in many edge-based AI applications [31], [32].

AI continues to penetrate new segments at a rapid pace. Currently, digital industries such as finance, retail, advertising, and multimedia are the sectors that have used AI the most. AI has added real value in these areas. Nevertheless, in several other areas, there are crucial problems that still need to be solved. The solution to cities' problems in transportation, energy and water supply, citizen safety, healthcare, and many others is to replace or improve old and ineffective technologies. New and AI-driven technologies have the potential to enable efficient transportation systems, clean energy, efficient healthcare systems, and efficient industry [33], [34], [35]. A critical element in these areas is the deployment and use of intelligence "at the edge" of high-speed and broadband networks [36]. Bringing intelligence to the edge signifies that even the tiny devices deployed

anywhere are able to sense, learn from, and respond to their surroundings. For instance, AI empowers devices situated on specific city streets or public spaces to make higher-level decisions, operate autonomously, and report noteworthy anomalies or enhancements to affected users or the cloud infrastructure.

Edge AI (or edge intelligence) means that AI algorithms are executed locally on a hardware edge device [37], [38]. The AI device can process data generated on the device and make decisions independently without requiring connectivity to function correctly. Using Edge AI necessitates from the device to have sensors connected to a small microcontroller unit (MCU). The MCU is loaded with specific machine-learning models trained for typical scenarios the device encounters. The learning process can also be continuous so that the device learns as it encounters new situations. The AI response can be a physical actuation in the device's immediate vicinity or a notification to a specific user or the cloud for further analysis and support.

GPU clusters reign supreme in the cloud-based machine learning and deep learning landscape, empowering AI workloads with unparalleled computational prowess. This type of specialized hardware for machine learning (ML) workloads is impractical for many edge resources due to their size and power requirements. Instead, specialized hardware has recently emerged to accelerate certain compute- or I/O-intensive operations at the network's edge. These edge hardware accelerators include Google's Edge Tensor Processing Unit (TPU) [39], [40], Nvidia's Jetson Nano and TX2 edge GPUs [41], [42], Intel's Movidius Vision Processing Unit (VPU) [43], and Apple's Neural Engine. They are explicitly designed for edge computing to support AI

applications such as visual and speech analytics, face recognition, object recognition, and deep learning inference.

Edge computing and edge AI encompass operations such as collecting, parsing, aggregating and routing data, as well as rich and advanced analytics that include machine learning, event processing, and actions at the edge. Edge AI will enable the execution of real-time operations, including data creation, decision-making, and response when milliseconds count. Real-time operations are essential for monitoring public places with crowds, self-driving cars, robots, monitoring machines in a factory, and many other areas. Edge AI will reduce data communication costs and thus power consumption, as edge devices process data locally and transmit fewer data to the cloud, improving battery life, which is extremely important for edge devices.

Smart cities are ideal for the deployment of edge computing and edge AI. Sensors and actuators deployed in various city locations and systems can receive commands based on decisions made locally without having to wait for decisions made elsewhere, far away. Cities leverage edge computing for real-time video surveillance, enabling prompt corrective actions to enhance safety and prevent accidents in streets, intersections, and buildings. They can also use it for lighting control, energy and power management, water consumption, and more, shortening end-to-end latency and reducing network congestion. By processing data generated by edge devices locally, smart city facilities can avoid the problem of streaming and storing large amounts of data in the cloud, which compromises data privacy and leaves them vulnerable to attacks.

III. EDGE AI-ENABLED VIDEO ANALYTICS
A. AI-ENABLED VIDEO ANALYTICS
In recent years, the field of video analytics, alternatively referred to as intelligent video analytics, has captured substantial attention and interest from both academia and industry. Notably, video analytics has revolutionized tasks that were previously reliant on human intervention, introducing automation and efficiency. By using AI with video analytics, a localized framework for intelligent video analytics emerges, enabling organizations to deploy video systems that autonomously detect spatial and temporal events directly at the network's edge. The applications of such video capabilities extend beyond security, encompassing a wide array of vital use cases.

Video analytics leverages AI algorithms to facilitate the autonomous identification of individuals, objects, or events. What sets AI software apart from conventional software is its iterative nature, wherein deploying the model in a production environment is merely one aspect of the overall process. Instead, the acquisition of pertinent data on a regular basis, the training and evaluation of the model, and the subsequent repetition of this cycle are imperative to attain the desired level of accuracy. AI-driven video analytics can be augmented by two primary categories of AI algorithms: Machine Learning and Deep Learning (see Figure 2).
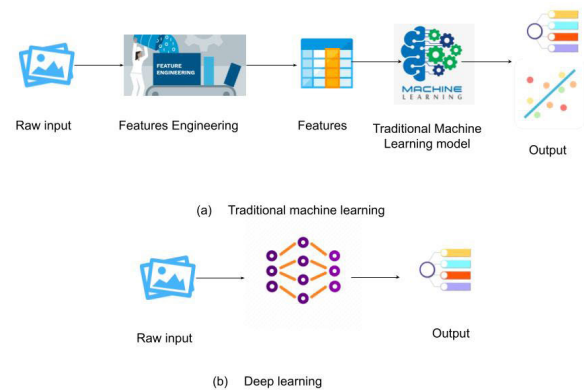

FIGURE 2. Machine learning and Deep learning processes.

1) MACHINE LEARNING
The machine learning process defines a data science team's workflow to build and deliver a machine learning model. It also defines how the team should work and collaborate to create the most useful predictive model possible. When using machine learning for video analytics, raw data is collected from various sources, recorded, and labeled based on features. After an iterative training process, a fully trained traditional machine learning model yields a program that can be used for output in the form of event and image recognition and identification. As the algorithm continues to perform its function in production with the given data, it learns and improves over time. Traditional machine learning methodologies rely on fixed and manually crafted features rather than automatically learned features derived from the available data, limiting their potential for achieving optimal performance and accuracy across various scenarios.

2) DEEP LEARNING
Deep learning, widely regarded as a subdivision of machine learning, encompasses various algorithms that hinge upon neural networks as their fundamental architectural framework. The recent prominence of neural networks stems from the convergence of two transformative factors: the abundance of Big Data and the accessibility of cost-effective parallel computing hardware like GPUs and computer clusters. Employing computational intelligence, deep learning is a preeminent paradigm for knowledge acquisition, experiential learning, and developing intricate concepts from simple ones.

Deep learning differs from traditional machine learning by being able to learn features from input data autonomously. Through the training process, a Deep Neural Network analyzes thousands of labeled images to develop its classification capabilities. When the algorithm is shown an input image, the neural network layers respond to increasingly complex shapes and structures and compare them to the training data. Finally, an identification of the image is made. The same is true for video data. Deep learning possesses the remarkable capability to autonomously identify and extract relevant

features, even within concatenated sets, thereby empowering it to adeptly recognize and classify complex objects.

B. EDGE VIDEO ANALYTICS

The contemporary business landscape has witnessed an intensified reliance on data, highlighting the significance and intricacy of information technology (IT) architecture. Present-day enterprises increasingly adopt a data-driven approach, leveraging the proliferation of Big Data and IoT devices. Consequently, these organizations confront an unprecedented surge in the data points utilized within a given application. Integrating data obtained from diverse endpoints, such as IoT sensors and cameras, engenders a proportional escalation in the magnitude of data necessitating processing, storage, and management. Furthermore, the advent of AI applications and the deployment of high-resolution cameras have engendered a notable upsurge in data accumulation within video analytics.

Cloud-based data processing, though pervasive, can suffer from bandwidth bottlenecks and delays, especially in critical applications like security. Edge AI comes to the rescue in many of these use cases. Edge video analytics offers several advantages. For example, bandwidth issues can be avoided, delays in data access can be reduced, and compliance rules and regulations can be met. Nevertheless, the most crucial benefit is that video analytics with Edge AI provides faster on-site insights and makes critical decisions in real-time.

Of course, video analytics can be performed via centralized servers or the cloud. Storage and execution of commands derived from video analytics have only recently been introduced in edge computing and edge AI. It is predicted that 64% of all IP cameras shipped worldwide will be AI-enabled cameras, providing the opportunity to perform much of the AI work at the edge. As AI moves into the mainstream market, intelligent video analytics is likely to be integrated into many cameras due to its impressive capabilities.

Combining Edge AI and video analytics holds immense practical advantages, poised to revolutionize the security industry. AI represents the imminent progression of the CCTV camera market, enabling the automation of camera footage or live image processing and analysis. Notably, security stands as the primary domain poised to reap substantial benefits from the convergence of Edge AI and Video Analytics, primarily due to the accelerated on-premises data processing capabilities, surpassing the efficiency of cloud-based counterparts. Furthermore, integrating Edge AI and Video Analytics will strengthen search and monitoring functionalities. The automatic identification of license plates, faces, and pedestrians and adherence to workplace safety protocols can be accomplished with greater precision and practicality, facilitated by the prowess of Edge AI and video analytics.

In addition, Edge AI video analytics data processing is done on-site with machines nearby. Let us say something needs to be fixed with security or quality control (another important use case for video analytics); the last thing a company wants is for its video analytics algorithm to struggle with latency issues that can occur with the cloud. When AI and data processing are done on-premises, the time it takes for the video analytics algorithm to send an alert is shortened, leaving more time to fix the detected security breach or quality issue.

C. EDGE AI PLATFORMS FOR VIDEO ANALYTICS

Edge AI platforms redefine the computational landscape by executing AI models at the network edge rather than in centralized data centers or the cloud. Edge AI empowers real-time analysis of locally generated data, eliminating the need for data transmission to remote locations. This scenario benefits video analytics applications by allowing immediate analysis and response to events captured on video data. Edge AI offers the invaluable advantage of minimizing data transmission to cloud servers or data centers, streamlining computational efficiency. It can benefit applications operating in low-bandwidth or offline environments or applications that generate vast volumes of data, such as high-resolution video streams. Nevertheless, Edge AI also has drawbacks. One of the main challenges is the limited computational resources of Edge AI devices, making it difficult to run complex AI models, which can then limit the video analytics application's capabilities and the results' accuracy. Besides, the cost and complexity of deploying the application on Edge AI devices can increase when these devices require specialized hardware and software. Edge AI platforms thrive in real-time video analysis applications like traffic monitoring, security, surveillance, and industrial automation. They are also helpful for applications working in low-bandwidth or offline environments. However, a cloud-based AI solution is better suited for more complex video analysis requiring vast amounts of data or high accuracy.

Several Edge AI platforms and open-source software are available for video analytics, including NVIDIA Jetson [41], [42], Google Edge TPU [40], Intel OpenVINO [43], Qualcomm Neural Processing SDK [44], and OpenCV [45]. Each of these Edge AI platforms has its benefits and drawbacks. Choosing the ideal platform for a particular application necessitates meticulous consideration of the unique requirements inherent to a given application.

IV. METHODOLOGY

This review article employs qualitative research methodology to synthesize relevant literature on video analytics, fostering a comprehensive understanding of the subject. In light of the inherent descriptive nature of this study, adopting a qualitative approach is essential to reviewing and merging a large body of relevant scientific literature. Without aiming for complete coverage, a meticulous and systematic examination strategy has been adopted to achieve the objective of this undertaking.

A. SEARCH CRITERIA FORMULATION

The search criteria used were:

- C1: ("Edge AI" OR "edge intelligence") AND "Video Analytics"
- C2: ("AI" OR "edge intelligence") AND "Video Analytics"
- C3: ("AI" OR "artificial intelligence") AND "Video Analytics"
- C4: "Edge Video Analytics"
- C5: "Deep learning" AND "Video Analytics"
- C6: "Machine learning" AND "Video Analytics"
- C7: ("smart" OR "intelligent") AND "Video Analytics"
- C8: "Privacy" AND "Video Analytics"

The purpose of this review paper is to answer the following research questions.

- **RQ-1**: What are the myriad applications of Edge AI and video analytics in the contemporary smart city landscape? This research question strives to uncover cutting-edge efforts and breakthroughs in the use of Edge AI and Video Analytics technology in key areas within a smart city.
- **RQ-2**: What machine learning and Deep learning models are used in edge video analytics?
- **RQ-3**: What techniques and methods are used for privacy-preserving in edge video analytics?
- **RQ-4**: What are the potential open research issues and future directions in Edge AI and Video Analytics implementations in a smart city? This question seeks to define the unanswered inquiries and unexplored paths that hold the key to unlocking the full potential of Edge AI and Video Analytics within the context of smart cities. By unraveling the challenges that may hinder their widespread adoption and delving into research directions, this query drives researchers to understand the current landscape of edge intelligence and Video Analytics, unraveling novel insights and paving the way for transformative advancements in this domain.

B. SOURCE SELECTION AND APPROACH

An extensive exploration was undertaken utilizing various authoritative databases and search engines to amass pertinent research material for this review. Four popular databases (Scopus, IEEE Xplore, Web of Science, and Google Scholar) renowned for their comprehensive coverage were used for the search of scholarly works on the subject. The search strategy revolved around targeted search criteria, focusing on the key terms "Edge AI" and "Video Analytics" while augmenting the search scope with synonymous terminologies such as "edge intelligence" and "edge computing" to broaden the search outcomes. A time constraint was applied, confining the search to encompass articles published between 2018 and 2023.

Most of the papers reviewed are Journal articles or conference papers. They were selected based on journal quality and relevance to the topic and filtered by The articles' selection is based on titles relevant to the topic of this review. References published before 2017 cited in this review paper mainly con-

TABLE 1. Distribution of the collected references.

Topic	N. of references
VA at the Edge	55
VA using deep learning	44
VA using Machine Learning	41
Intelligent video analysis	36
Federated learning for VA	7
Edge cloud collaboration for VA	21
VA for Surveillance and Security	15
Privacy-preserving VA	7
VA for Transportation	12
VA for Medical Applications	11
VA for Sports and Entertainment	12
VA for Education	7
Intelligent VA for Various Applications	14

cern the background and literature review on edge computing and video analytics.

V. RESULTS

The initial search for the above search criteria (C1 - C8) found 109 references from Scopus, 152 from Web of Science, 85 from IEEE Xplore, and 406 from Google Scholar. However, the total number of references, 752, was reduced to 408 after eliminating duplicates. Further screening permitted the elimination of 126 references addressing issues far from the main topic of this review paper. Screening the abstract and full text permitted the elimination of 170 more references, mainly about video analytics done in environments other than at the edge. The final number of references eligible for this study is 112. These references do not include the references we cited to provide background information on Edge computing, Edge AI, and Video Analytics.

Figure 3 shows the PRISMA diagram that represents the different phases of this systematic review process [46]. After eliminating duplicate references from the four bibliographical databases used in this study, the analysis of the titles of the 282 references identified permitted to draw a classification of the different topics addressed in these references.

Table 1 shows the number of references found in each category (or topic). Figure 4 shows the distribution of the 282 collected references, and Figure 5 depicts the distribution of the 112 eligible references obtained in the last phase of the selection process, which consists in eliminating references that do not use edge computing after screening their abstracts and full text. The analysis of the abstracts of the eligible works of this study permits to draw in Figure 6 the taxonomy of the various topics addressed in these works. The next subsections describe the findings concerning each class of this taxonomy.

A. VIDEO ANALYTICS AT THE EDGE

The analysis of the works on Video Analytics at the edge from our final list of eligible articles resulted in their classification into two subclasses: Edge-based real-time video analytics and Optimization and efficiency of edge-based video analytics as depicted in Figure 6.
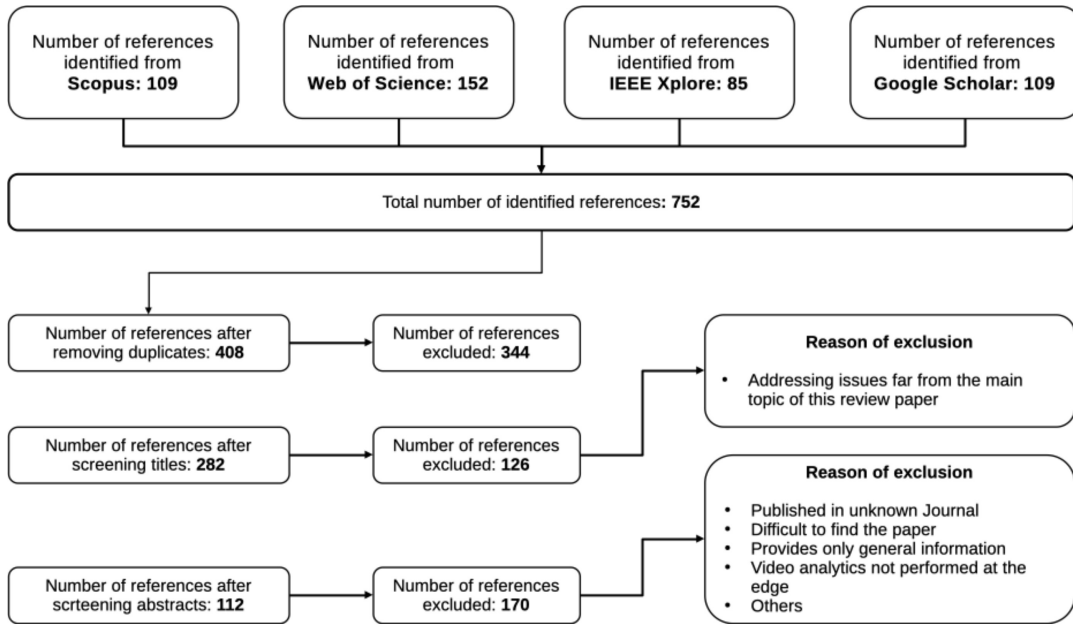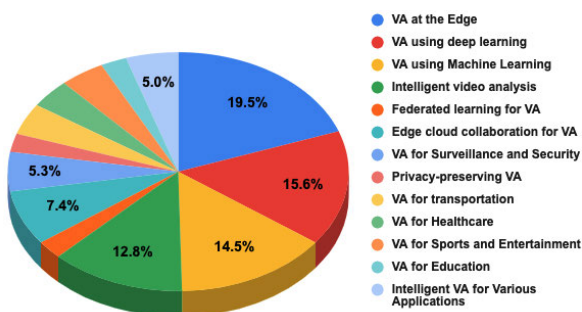
FIGURE 3. Prisma diagram of the systematic review.

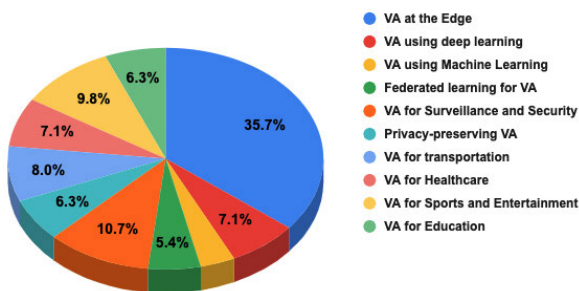**FIGURE 4.** Distribution of the collected references.

FIGURE 5. Distribution of the eligible references of this study.

1) EDGE-BASED REAL-TIME VIDEO ANALYTICS

Further analysis of the works in this class permits us to classify them into three categories:

- Efficient video analytics at the edge
- Advanced techniques for edge video analytics
- Applications of edge video analytics

Table 2 and Table 3 summarize the various works in the first and second categories. The third category, which is about

the various applications of edge video analytics, is described in detail in Section VI.

2) OPTIMIZATION AND EFFICIENCY OF EDGE-BASED VIDEO ANALYTICS

Further analysis of the works in this class permits classifying them into three categories:

- Scalable video analytics at the edge
- Ege video analytics for specific applications
- Data-efficient and opportunistic edge video analytics

Table 4 and Table 5 summarize the various works in the first and third categories.

B. VIDEO ANALYTICS LEARNING MODELS

This section addresses **RQ-2**. The review of the final list of eligible references of this study shows that several machine learning and deep learning models are used in edge video analytics for various tasks such as object detection, tracking, recognition, and classification. Besides, several works used federated learning for the training of their models. Table 2, Table 3, Table 4, Table 5, Table 6, and Table 7 describe the various models used in the works of each of the above categories. The commonly used models are:

- *Convolutional Neural Networks (CNN)*. CNNs are pivotal in edge video analytics, excelling in object detection and recognition. Their capacity to handle substantial quantities of visual data and discern complicated patterns is of utmost importance in the interpretation of meaningful observations from visual information.
- *Recurrent Neural Networks (RNNs)*. RNNs serve a pivotal function in the realm of video analysis by effectively managing tasks such as tracking, segmentation, and
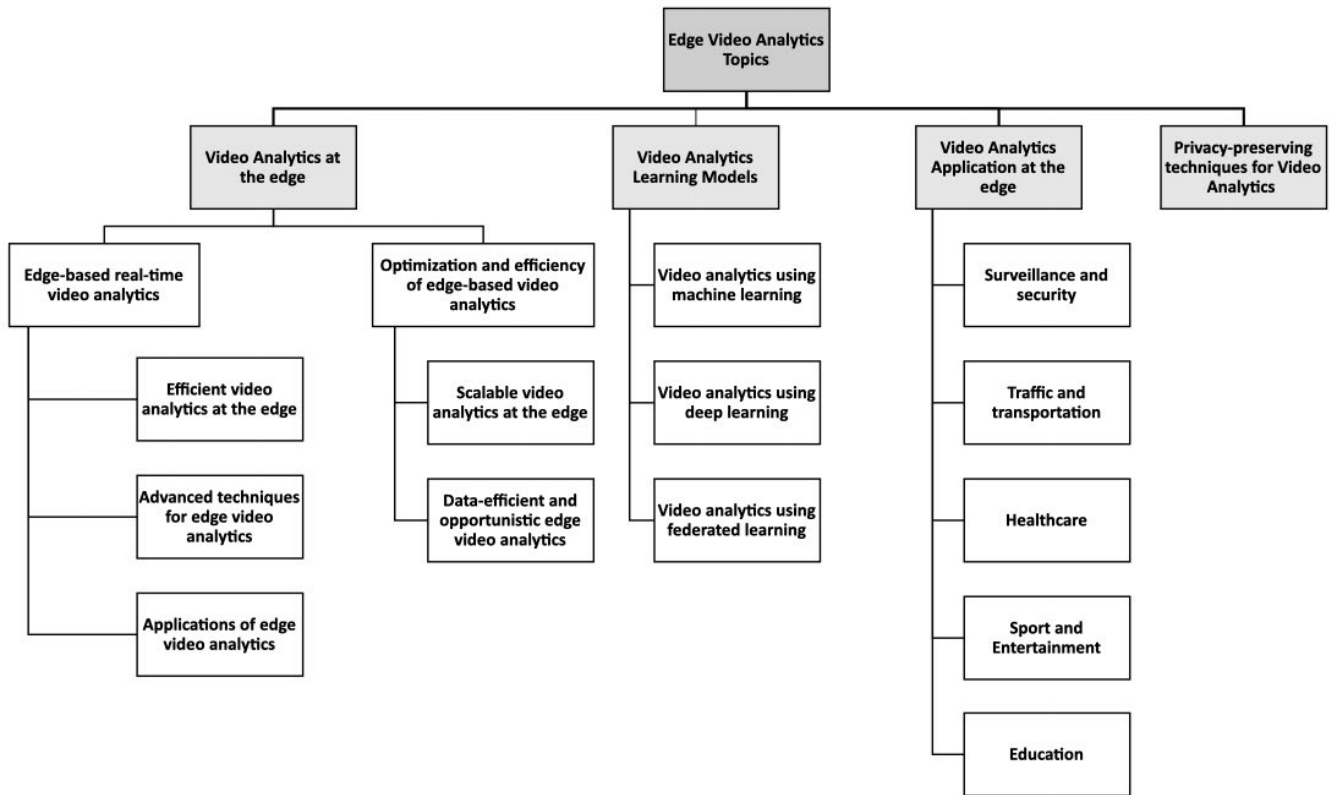
FIGURE 6. Taxonomy of the eligible references of this review.

action recognition. With their proficiency in processing sequential data, RNNs are well-suited for analyzing videos, essentially temporal image sequences.

- *You Only Look Once (YOLO)*. The YOLO model represents a noteworthy real-time object detection model that has proven to be highly effective in edge video analytics. Its inherent capacity for detecting objects in a single pass renders it a more expeditious alternative to other object detection models [74], [75].

- *Mask R-CNN*. It is a profound neural network architecture employed for the purpose of detecting objects and segmenting tasks in the domain of edge video analytics. Notably, it possesses the ability to effectively detect objects and yield pixel-level masks for every object present in an image [76].

- *MobileNet*. The utilization of the MobileNet deep learning framework in edge-based video analytics is a common practice deployed for diverse applications, including object detection and recognition, primarily due to its minimalistic design. The fundamental aim of this framework is to operate with optimal efficiency on mobile devices that possess restricted processing capabilities [77].

- *Single Shot Detector (SSD)*. The SSD model presents a real-time remedy for the identification of objects within edge video analytics. Its singular capacity to execute object detection in a solitary pass elevates its efficiency beyond that of other models [78].

- *Inception*: Inception, a powerful deep learning model, finds its application in edge video analytics for various tasks, including image classification and object detection. With its optimized architecture, Inception excels at efficiently processing high-resolution images, enhancing the overall performance of edge video analytics systems [79].

Various metrics and benchmarks are commonly employed to assess the efficacy of Edge AI algorithms for video analytics. The key metrics of utmost significance comprise the following:

- *Accuracy*. The evaluation of Edge AI algorithms for video analytics is commonly conducted by utilizing accuracy as the primary metric. This metric serves to gauge the algorithm's efficacy in detecting objects or events in a given video stream.

- *Precision and Recall*. Precision and recall are deemed as essential evaluation metrics for algorithms aimed at detecting objects. Precision quantifies the ratio of true positives to false positives, scrutinizing the algorithm's precision in correctly identifying objects. In parallel, recall captures the ratio of true positives to false negatives, shedding light on the algorithm's capability to avoid missing objects of interest. These metrics reveal the algorithm's efficacy and proficiency in object detection endeavors.

- *Intersection over Union (IoU)*. IoU is a crucial metric for evaluating object detection algorithms. It quantifies

TABLE 2. Summary of research on efficient video analytics at the edge.

Ref.	Year	Application	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[47]	2021	Low-cost edge-based real-time video analytics, with a focus on object detection and tracking in surveillance videos	Raspberry Pi 3 Model B+ as the edge device and a laptop as the server	Convolutional Neural Network (CNN)	the Pascal VOC 2007 and MS COCO datasets for training their object detection and tracking models	Approach achieved an accuracy of 82.8% and a processing time of 126 ms per frame on the Raspberry Pi 3 Model B+
[48]	2021	"Effective video analysis on mobile edge devices, with a focus on placing DNN models on edge devices to achieve better performance"	Raspberry Pi 4 Model B	Deep Neural Networks (DNNs)	The COCO dataset for object detection and the UCF-101 dataset for action recognition	Approach achieved 91.8% accuracy and 104 ms latency for object detection, and 89.3% accuracy and 62 ms latency for action recognition on Raspberry Pi 4 Model B.
[49]	2021	Video analytics at the edge	Intel Xeon E5-2680 v3 CPU, NVIDIA Tesla K40 GPU	YOLOv3 (You Only Look Once version 3)	PASCAL VOC, MS COCO	Proposed approach achieved 3x faster processing time than traditional sequential approach with high accuracy for object detection.
[50]	2022	Deep video analytics at the edge	NVIDIA Jetson TX2, Raspberry Pi 4 Model B	Residual Neural Network (ResNet)	UCSD Ped1, ShanghaiTech, VIRAT	Proposed approach achieved comparable accuracy to traditional approach with up to 50% reduced detection time according to evaluation results.
[51]	2022	Live Video Analytics at the Edge	Intel i5 processor with 16GB RAM	Multiple models, including Faster R-CNN and Mask R-CNN	Not mentioned	VPPlus achieves over 95% high accuracy level for live video analytics at the edge with significantly lower processing time (up to 10x faster) compared to state-of-the-art method.
[52]	2021	Low-cost edge video analytics system with velocity-based configuration adaptation.	Raspberry Pi 4 Model B and a camera module	YOLOv3, object detection algorithm	2000 images of vehicles, pedestrians, and bicycles.	Reduced false alarms by 24.2% and improved processing time by 31.7% compared to the baseline system.
[53]	2022	Analyzing real-time video streams from moving vehicles	Not specified	YOLOv3	The KITTI dataset	System achieved 92.4% average precision with potential applications in autonomous driving, traffic monitoring, and surveillance.
[54]	2021	Methodology to evaluate and compare the performance of different edge video analytics systems.	N/A	N/A	N/A	Paper provides guidelines for designing efficient edge video analytics systems.
[55]	2021	Low-latency video analytics	Not specified	Reinforcement learning-based algorithm for resource allocation and query scheduling	Real-world video analytics dataset	Proposed edge learning framework achieved 30% mean inference latency reduction compared to baseline method.
[56]	2018	Live video analytics for drones	Raspberry Pi 3, NVIDIA Jetson TX2, and DJI Phantom 4 drone	Combination of YOLO object detection and KCF object tracking algorithm.	PASCAL VOC and ImageNet datasets	System achieved real-time performance for drone video analytics with 95% detection rate, 2.3 fps on Raspberry Pi, and 29.3 fps on NVIDIA Jetson TX2. Reduced bandwidth usage by 80% through edge computing.

the overlap between predicted and ground truth bounding boxes, assessing algorithm precision in identifying objects.

- *Frame rate*. The frame rate metric is paramount as a key criterion to assess the algorithm's performance, quantifying the rate at which images are processed within a unit of time. In real-time applications demanding an increased frame rate, ensuring the algorithm stays in sync with the streaming video's continuous flow becomes indispensable.

- *Latency*. This metric measures the time an algorithm takes to process a frame and deliver a result. Low latency

is critical in real-time applications, as it empowers algorithms to present results quickly, leaving no room for delays.

- *Power consumption*. With the emergence of Edge AI as a focal point, the imperative of governing power consumption assumes a paramount role. Indeed, optimizing power consumption prolongs battery life and facilitates cost reduction, thereby underlining its criticality in this domain.

- *Memory usage*. Given the constraints imposed by limited resources, considering memory consumption assumes a position of utmost importance in Edge AI.

TABLE 3. Summary of research on advanced techniques for edge video analytics.

Ref.	Year	Application	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[57]	2022	Integrating quality control into the video analytics process.	Not specified	ResNet-based human activity recognition CNN	Experiments on an edge computing platform	Improved video quality and reduced processing delay.
[58]	2022	Adaptive cross-camera video analytics	Not specified	Spatial and temporal correlation models in cross-camera video analytics	Market-1501 dataset	Comparison with other benchmarks considers Computational cost, Recall and precision, Heavy hitter analysis, and Sensitivity of time granularity.
[59]	2022	Edge AI Real-Time Video Analytics	Analog Matrix Processor	Deep Neural Networks (DNNs)	COCO dataset	High performance and energy efficiency compared to digital alternatives. Pose detector metrics: latency, FPS, Power, mAP.
[11]	2022	A survey on the state-of-the-art in deep learning-driven edge video analytics	N/A	N/A	N/A	Paper comprehensively covers current state-of-the-art, including challenges and opportunities.
[60]	2020	Using natural language queries to perform video analytics at the edge	NVIDIA Jetson Nano	SENet-154 model pre-trained with ImageNet	ActivityNet Captions and MSRVTT datasets	Outperforms baseline methods in terms of F1 score, bandwidth consumption, and runtime efficiency.
[61]	2021	Smart distribution and processing of vision modules in parallel across available edge compute nodes for live edge video analytics	NVIDIA Jetson Nano	YOLOv4 and OpenCV	Not mentioned	Evalutes the change in accuracy of bounding boxes with change in input resolution and the change in Memory utilization of blocks.

Exercising due caution becomes imperative as it minimizes memory usage, guaranteeing a seamless operation on devices with limited resources at their disposal.

C. PRIVACY-PRESERVING IN EDGE VIDEO ANALYTICS

The present subsection explores **RQ-3**, which pertains to privacy-preserving techniques and methods within edge video analytics. These methods play a vital role in safeguarding personal information while concurrently facilitating video data analysis. A summary of the different scholarly works examined in this analysis, with a particular emphasis on the employment of privacy-enhancing methods in edge video analytics, is delineated in Table 7. The common techniques and methods used for privacy-preserving in edge video analytics include:

- *Encryption*. Encryption involves the methodical conversion of data into an encrypted format using cryptographic algorithms, making it incomprehensible without the corresponding decryption key. This powerful technique finds application in safeguarding sensitive information, including but not limited to biometric data, during its transmission or storage, thereby ensuring an elevated level of security and confidentiality.
- *Differential privacy*. Differential privacy, an ingenious technique, introduces controlled perturbations to data with the dual purpose of safeguarding individual privacy while simultaneously enabling the extraction of valuable insights. Its application in edge video analytics proves particularly advantageous in preserving the anonymity and identities of individuals captured in video data, thus reinforcing the protection of their privacy.
- *Anonymization*. Anonymization, an integral process, entails the meticulous eradication of personally identifiable information from data, thereby assuring the

concealment of individuals' identities. When judiciously applied, this technique serves as a valuable means to protect the privacy of individuals featured within video data while concurrently enabling comprehensive analysis of the underlying data.
- *Edge computing*. Edge computing entails the localized processing and analysis of data at the network edge, in close proximity to its origin, enabling swift and efficient decision-making. This reduces data transmission, minimizing the risk of data breaches and unauthorized access.
- *Federated learning*. Federated learning represents an intriguing machine learning paradigm, wherein models undergo training leveraging distributed data from numerous devices or locations, all while circumventing the necessity of data migration towards a central server. Such a technique holds the potential to ensure the preservation of individuals' privacy within video data by upholding its decentralized nature, thereby amplifying the overall sense of perplexity and fostering a robust privacy framework.
- *Multi-party computation*. Multi-party computation (MPC) is a technique that ensures the collective computation of a function while safeguarding individual input confidentiality. It fortifies privacy in edge video analytics, enabling insightful analyses. By preserving individuals' privacy, MPC reinforces security and facilitates invaluable analyses.

VI. OPPORTUNITIES AND USE CASES OF EDGE AI-ENABLED VIDEO ANALYTICS

This section addresses **RQ-1**. Gartner predicts that 50% of all inferences will take place at the edge by 2025. It classifies edge video analytics as being in a youthful stage of maturity

TABLE 4. Summary of research on scalable video analytics at the edge.

Ref.	Year	Application	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[62]	2020	Scalable Video Analytics	Intel Movidius Neural Compute Stick (NCS) and Raspberry Pi 3 Model B+	Compressed-Domain Convolutional Neural Network (CD-CNN)	WIDER FACE and PASCAL VOC datasets for CD-CNN model training, along with a custom dataset created from compressed video frames.	Higher accuracy (avg. precision of 0.87) than prior methods on the WIDER FACE dataset with CD-CNN model. Achieved scalability and near-real-time performance on Raspberry Pi 3 Model B+.
[63]	2020	Live Video Analytics	NVIDIA Jetson AGX Xavier and Jetson Nano	Custom DL model combining CNNs and LSTM networks.	COCO dataset and custom traffic video dataset for DL model training.	Achieved 89.5% accuracy on the COCO dataset for object detection and 92.5% on the custom traffic dataset for vehicle detection. Real-time processing with significantly lower latency and communication overhead compared to traditional cloud-based methods, up to 4x and 9x, respectively.
[64]	2023	Real-Time Video Analytics for IoT	A custom edge device based on the NVIDIA Jetson TX2 and a Raspberry Pi 4	Deep learning framework called OsmoticNet, which combines a lightweight CNN model with a long short-term memory (LSTM) model.	ImageNet dataset for the pre-training model and a custom dataset of traffic videos for fine-tuning and evaluating the model.	Achieved up to 2.4x and 4.3x lower latency than traditional cloud-based and edge-based methods while maintaining comparable accuracy.
[65]	2020	Low-latency and accurate video analytics	Not specified	CNN classifier	RetinaNet-50	The median accuracy achieved by the proposed algorithm increases from 64%,73%, to 76% when the edge computing resource increases from 50, 100, to 300 GigaFlop/s respectively.
[66]	2022	Efficient edge-based industrial video analytics	Edge Nodes: i) CPU: Intel(R) Core(TM) i7-9750H CPU, 8 GB RAM. ii) AMD Ryzen 7 4800H with Radeon graphics, 8 GB RAM. Cloud server: Intel(R) Core(TM) i9-10900K 32 GB RAM.	Not mentioned	EPFL, a multicamera pedestrian video dataset, VIRAT dataset, a high-resolution video surveillance dataset, and PETS09 dataset	Used the following performance metrics: Computational Cost, Accuracy, Reduction Ratio, Latency, and Network bandwidth.
[67]	2022	Configuration selection and bandwidth allocation for multiple video streams in Edge-based video analytics	Not specified	Different CNN models are deployed on the single edge server to match various video resolutions	Not specified	Results demonstrate that the proposed solution effectively balance analytics accuracy, energy consumption, and system latency across various settings.
[68]	2022	Modeling multi-cell edge video analytics	Not specified	YOLOv5 and a mathematical model for optimizing the performance of edge video analytics systems across multiple cells	Not specified, simulation experiments	Analyzed impact of system parameters on model's performance. Developed a promising mathematical model for multi-cell edge video analytics to optimize system performance.
[69]	2019	Scalable video analytics at the edge	HP Z840 workstation with 32 GB RAM and an Intel Xeon processor with 10 physical (20 virtual) cores and a dedicated NVidia Titan XP GPU with 12 GB RAM.	Faster-RCNN	Duke MTMC dataset, VOC 2012 dataset	The proposed system shows average throughput gains of more than 4x by trading off a 5% drop in inference accuracy.
[70]	2018	Large-scale video stream analytics at the edge	Cloud cluster with 8 compute cloud nodes, Each with 100 GB, 4 VCPUs and 16 GB RAM. Configuration of edge nodes not mentioned.	Convolutional Neural Networks (CNNs)	The total video dataset size is 5GB	Results showed efficiency gain of 37% and 71% in Edge-Cloudlet-Cloud and Edge-CloudletCloud-Filter configurations, respectively, when compared to the cloud-only approach.

with transformative potential. The advantages of integrating edge AI and video analytics are widely recognized, leading

experts to forecast a growing trend of video analytics applications migrating toward the edge.

TABLE 5. Summary of research on data-efficient and opportunistic edge video analytics.

Ref.	Year	Application	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[71]	2023	Data-efficient continuous learning for edge video analytics	Real-world vehicle cameras	ResNet18, MobileNet-V3, and Faster RCNN	CORe50 dataset, and real-world dataset of DashCam	The proposed system can save 70% labeling and storage costs while maintaining a competitive accuracy. It achieves up to 47.84% AP accuracy on the real-world DashCam dataset.
[72]	2022	Edge video analytics that leverages idle resources of edge nodes	Azure Stack Edge Pro equipped with an NVIDIA Tesla T4 GPU, and a virtual machine equipped with an NVIDIA Tesla V100 GPU	EfficientDet-D0, Faster-RCNN, YOLOv3	UA-DETRAC dataset, AICity dataset, BDD100K dataset, and COCO dataset	The proposed system improves the absolute mAP by 9.02%, 11.34%, and 7.27% on average for EfficientDet-D0, YOLOv3, and Faster RCNN, respectively.
[73]	2022	Real-time video analytics that uses spatial-temporal redundancy suppression to reduce data transmission	Server with an octa-core Intel processor 3.3 GHz with 16 GB of RAM and NVIDIA GeForce GTX3060 graphics card with 6 GB of RAM, and Nvidia Jetson Intelligent Toy Car	Tiny YOLO, Nvidia TensorRT, Faster-RCNN, ResNet101, Deep neural network	Not specified	Proposed system outperforms several baseline systems. It reduces bandwidth usage by up to 90% while meeting the desired accuracy.

Edge AI and video analytics are a powerful amalgamation that is poised to redefine the realm of security across diverse business landscapes, ranging from large corporations to local enterprises. Additionally, traffic monitoring, retail analytics, quality control, and recognition tasks are all set to benefit from the integration of Edge AI in video analytics. With its numerous advantages, edge systems for video analytics allow for more diverse and efficient applications.

There are many different applications of edge video analytics, and they can be grouped into several categories or taxonomies. As depicted in Figure 7, an example of taxonomy for edge video analytics applications includes:

- *Surveillance and security*. The convergence of video surveillance, facial recognition, object detection and tracking, and license plate recognition engenders a multifaceted ecosystem meticulously devised to enhance security measures while vigilantly scrutinizing the surroundings for plausible threats.
- *Retail and marketing*. The deployment of diverse applications such as people counting, customer behavior analysis, and queue management within retail environments encapsulate a rich arsenal meticulously curated to enhance the customer experience while strengthening sales revenue.
- *Traffic and transportation*. An array of cutting-edge applications, including traffic flow analysis, license plate recognition, and object detection and tracking, are vital in the relentless pursuit of optimizing traffic dynamics and bolstering road safety measures.
- *Industrial and manufacturing*. Various applications, including quality control, process monitoring, and equipment maintenance, revolutionize industrial and manufacturing environments. This diverse array

collaboratively fosters enhanced operational efficiency, decreased downtime, and unrivaled product quality.
- *Healthcare*. Applications such as patient monitoring, fall detection, and remote assistance revolutionize healthcare facilities, enhancing patient care, operational efficiency, and cost reduction.
- *Environment monitoring*. Using cutting-edge technologies, cities can observe wildlife populations vigilantly, assess air quality meticulously, and decipher meteorological patterns.
- *Sports and entertainment*. The deployment of diverse edge computing-based applications, encompassing crowd management, audience engagement, and event analysis, propels a wave of transformations within sports and entertainment venues, allowing enhanced fan experience and augmented revenue streams.
- *Education*. Video analytics in education permits automated student assessment, personalized content delivery, behavior monitoring, adaptive learning, increased safety, and improved teaching effectiveness and student engagement.

The field of edge video analytics is steadily advancing at an unprecedented pace, leading to the proliferation of innovative applications across diverse industries.

A. SURVEILLANCE AND SECURITY

The utilization of edge video analytics significantly improves the efficiency of surveillance and security systems by facilitating prompt analysis and automated identification of questionable activities or occurrences. Several studies have shown the benefits of using edge computing and video analytics to ensure the safety of citizens in public places from crime, theft, and violence. As depicted in Figure 8, use cases of edge video analytics in surveillance and security include intrusion

TABLE 6. Summary of research on edge video analytics using deep learning and federated learning.

Ref.	Year	Application	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[80]	2018	Edge video analytics	Backend server Quad-core Intel processor (2.7 GHz, 8GB RAM) NVIDIA GeForce GTX970 4GB RAM graphics card, and Samsung Galaxy S7 frontend device	CNNs, Android TensorFlow, Tiny-YOLO, Big-yolo deep learning model (22 convolutional layers).	The visual object classes (VOC) image dataset [81].	There are tradeoffs in accuracy, battery consumption, bit rate, and data usage based on system characteristics such as model size, offloading decision, video resolution, and network parameters.
[82]	2018	Surveillance. Real-time object and face detection, recognition, and analytics.	Not mentioned	- CNN models: AlexNet, GoogLeNet (Human Sub-Attribute Recognition), YOLO and Tiny YOLO models, Pre-trained CNNs, Deep Face, DLib-ResNet, OpenFace, ResNet-SSD	Classroom Dataset (736 images), Community Center Dataset (5336 images), Traffic Dataset (999 images), Multi-Scenario Surveillance Dataset (8191 images), PETA Dataset, MNIST dataset, FERET Dataset.	Accuracy results for Human Sub-Attribute Recognition: Age: 77.4 % (AlexNet CNN model - Fine Tuning approach), Gender: 83.8 % (AlexNet CNN model - Fine Tuning approach), Apparel Color: 65.24 % (AlexNet CNN model Fine Tuning approach), Apparel Type: 64.1 % (GoogLeNet CNN model - Transfer Learning approach)
[83]	2023	Video analytics using deep learning on mobile devices through edge computing and neural processing units (NPUs).	HUAWEI 10 pro (6GB, octa-core CPU), NPU, HUAWEI DDK.	DNN models: VGG, AlexNet, YOLO Small, ResNet-152	LFW dataset, VOC dataset, MS COCO Dataset, FCVID dataset, ImageNet validation dataset.	VocNet experiences a 30% decrease in accuracy, while AlexNet shows an 11% decrease in accuracy. The F1 score for YOLO Small decreases to 0.3.
[84]	2022	Collaborative video analytics using multiagent deep reinforcement learning on distributed edges	Not mentioned	Faster-RCNN Mobilenet 320, Faster-RCNN Mobilenet, Retinanet Resnet-50, Mask-RCNN Resnet-50	Real-world datasets	The proposed approach significantly improves overall rewards from 33.6% to 86.4% compared to competitive methods (IPPO, Local-PPO, Shortest-Queue, Random, and Local).)
[85]	2019	Live video analytics using FPGA-based smart cameras	FPGA-based smart cameras, Intel Core i7 (3770 CPU), NVIDIA RTX 2080 GPU	The Gaussian Mixture Model (GMM), the Gaussian Distribution Models (GDMs)	Videos and images taken with a camera video (360x480 resolution) at a traffic intersection.	The proposed method outperforms CPU and GPU solutions in processing latency and energy efficiency. It achieved a 49x speedup in frames processed per second and a 6.1x improvement in energy efficiency compared to the CPU.
[86]	2018	Real-time intelligent video analytics using an edge service framework	A desktop (Intel i7-6700 CPU @ 3.40GHz, a Nvidia GeForce GTX 1060 6GB and 24GB system RAM) with Ubuntu 16.04 LTS.	Deep Neural Networks (DNNs): DetectNet neural network, YOLO, SSD	Not mentioned	High-performance and optimized inference using GPU acceleration and TensorRT.
[87]	2022	Context-aware video content analysis using deep learning on Internet of Things (IoT) devices.	A Nvidia Jetson Nano, a 128-core Maxwell GPU, a Quad-core ARM A57 CPU.	LSTM-based model, Transformer-based model, CNN, RNN.	MSR-VTT dataset	The transformer-based model achieves 97% accuracy, while the LSTM-based model enables low real-time inference.
[88]	2022	Task-oriented communication in the context of edge video analytics	Not mentioned	A temporal entropy model (TEM), MVDet, a ResNet-based model	The EPFL WILDTRACK Seven-Camera HD Dataset, the EPFL MVMC Detection dataset	The proposed TOCOM-TEM method consistently outperforms the baselines in terms of the rate-performance tradeoff.
[89]	2021	Use of AutoML (Automated Machine Learning) techniques for video analytics with edge computing.	A mobile device with a camera, a wireless 802.11ac Access Point, a server with a GPU for object recognition	Deep Neural Network (DNN) model, YOLO	COCO dataset	Stochastic accuracy and latency response have been observed even with fixed configuration parameters.

TABLE 6. *(Continued.)* Summary of research on edge video analytics using deep learning and federated learning.

[90]	2021	Collaborative cross-camera video analytics at the edge.	Cameras and an edge server	DNN model, YOLO	SALSA dataset, VOC2012 dataset	CONVINCE achieves around 91% accuracy in object identification while transmitting only about 25% of the recorded images.
[91]	2021	Federated video analytics architecture specifically designed for networked smart cameras.	Four Amazon DeepLens cameras, a server with an RTX-2080Ti GPU and an Intel i7 CPU 2.4 GHz	An LSTM model, A generic access network (GAN)-based model, Two CNN-based models	VeRi dataset (over 50,000 images of 776 vehicles), CityFlow (more than 10,000 images of 666 vehicles)	Ensures the privacy of video images and maximizes the accuracy of video analytics applications.
[92]	2019	A distributed video analytics architecture that leverages edge computing and federated learning techniques.	Not mentioned	Not mentioned	Images from road surveillance cameras: 960000 licensed vehicles, One million CCTV camera	Real-time Object detection and identification, Reduction in communication latency
[93]	2020	A federated video analytics system that utilizes edge computing	Dell XPS 8930 (Intel core i7-8700 12-Core Processor, Nvidia GeForce GTX 1080, 32GB RAM), Nvidia Jetson TX2	Neural Processes (NPs), Neural network (NN) model	COCO dataset, Video dataset for detecting objects and cars in video frames	By employing black-box and Neural Processes (NPs), the system's performance in terms of latency and accuracy is enhanced.
[94]	2019	A new edge computing architecture for distributed, scalable, and real-time video analytics based on Deep learning	Nvidia Xavier as the edge node, Nvidia Titan as the edge server.	TinyYOLOv3 for person detection/classification across many cameras, ShuffeNet-based network for generating feature vectors	Video Data, Frames of data transferred across the network, MPEG-4 and PNG images format	Object detection and re-identification on the edge server don't scale with increasing Edge node count, Object detection on the edge node and object re-identification on the edge server can be achieved by using a lightweight re-identification algorithm on the edge server. Local detection and re-identification on the edge node, along with global re-identification on the edge server, scale well with several edge nodes.

FIGURE 7. Use-cases of edge video analytics.

detection, object detection and tracking, facial recognition, abandoned object detection, crowd monitoring, and suspicious behavior detection.

1) INTRUSION DETECTION

Edge video analytics can detect unauthorized entry or intrusion into restricted areas. Analyzing real-time video streams,

TABLE 7. Summary of research on privacy-preserving techniques in edge video analytics.

Ref.	Year	Application	Privacy-preserving approach	Hardware	Learning model	Training datasets	Results (accuracy, ...)
[95]	2022	Person counting and Vehicle detection.	Edge side privacy control that leverages trajectory prediction and guarantees privacy leakage to a threshold.	Desktop with an Intel i9 CPU and an Nvidia RTX 3090 GPU, An AWS DeepLens camera	YOLOv5s and YOLOv5Face	nuScenes dataset, MOT15 dataset	Prevents privacy leakage under visual data attack with 95% accuracy and 4x speedup compared to existing privacy control mechanisms.
[96]	2022	Video analytics on cameras in public areas.	Event duration-based differential privacy	N/A	Faster-RCNN model in Detectron-v2 for object detection, and DeepSORT for object tracking.	Porto Taxi Service Trajectory - Prediction Challenge, ECML PKDD 2015 Dataset	Proposed system preserves privacy across various videos and queries with an error increase of 1-21% relative to a non-private system.
[97]	2021	Different applications of video analytics: (Generic Surveillance, Petty crime detection, people counting, ...)	Privacy assistants help configure privacy settings. Standardized APIs communicate opt-in/opt-out privacy decisions to users.	N/A	N/A	N/A	Diverse privacy preferences require advanced settings. Privacy assistants and standardized APIs are needed for user notifications and opt-in/opt-out communication.
[98]	2023	Identification of out-of-body scenes in endoscopic videos	Privacy preservation in surgical video analysis by identifying out-of-body images in endoscopic videos	N/A	A deep learning-based image classifier to identify out-of-body scenes in endoscopic videos (OoBNet).	A dataset created from surgeries recorded at the University Hospital of Strasbourg, Two independent multicentric hospital datasets used to validate the model.	Model achieved 99% ROC AUC in validation on 3 independent datasets.
[99]	2022	Identification of anomalous and dangerous situations involving violence or other harmful activities	Autoencoders and Differential Privacy	N/A	Double stream ResNet-50 network and YOLOv4 used for object detection. 4 optimizers: AdamOpt, DPSGDOpt, DPAdamOpt, DPadagOpt.	UCF-Crime dataset	System accurately detected video anomalies while maintaining privacy, shown by reduced face recognition in reconstructed frames.
[100]	2021	Mobile video analytics	Adversarial training extracts features for normal task classification and conceals sensitive task features to protect private information.	6 Mobile devices and embedded sensors	Convolutional neural networks (CNNs)	Affectiva-MIT Facial Expression Dataset (AM-FED+) and Privacy Evaluation Video Dataset (PEViD).	System protects private information with <2% accuracy drop while reducing execution time by up to 82.9% and energy consumption by up to 78.8%.
[101]	2022	Campus-wide video analytics applications	Only meta-data travels through the network to offer a privacy-compliant tracking solution.	Raspberry Pi 3 Model B+, NVIDIA Jetson Nano, Dell Precision workstations, NVIDIA Jetson TX2.	Not mentioned	Not mentioned	Study compares network bandwidth and storage consumption with and without CVAF but lacks specific performance metrics. Shows improved average response time during peak and non-peak hours.

the system can identify and raise an alert when a person or object crosses predefined boundaries or enters restricted zones. This helps security personnel to respond quickly and prevent potential security breaches. In the context of railroad

security, Zaman et al. in [102] implemented a video analytics and AI-based system to analyze railroad video cameras and detect unsafe track violations, thereby minimizing risks and potential accidents to residents.

2) OBJECT DETECTION AND TRACKING

Edge video analytics algorithms can detect and track objects of interest within a video feed. This includes identifying and tracking individuals, vehicles, or specific objects in real-time. The system is capable of activating alarms or issuing notifications upon detection of questionable objects or activities by consistently monitoring the video stream, enabling the implementation of proactive security measures. Zhang et al. [103] implemented an alert assistant AMBER based on data analysis from city cameras to monitor suspicious vehicles and help citizens find missing children or stolen vehicles.

3) FACIAL RECOGNITION

With edge video analytics, cameras can perform facial recognition at the edge without relying on centralized servers. This enables quick identification of individuals from a database of known persons or persons of interest. Facial recognition at the edge enhances security by identifying potential threats or identifying unauthorized individuals in real-time [9]. The authors in [104] developed an intelligent indoor video surveillance system based on a CNN deep learning algorithm to analyze and track the content of video cameras to provide security in prisons.

4) ABANDONED OBJECT DETECTION

Edge video analytics can detect and raise alerts when an object is left unattended or abandoned for a certain period. Unattended bags or packages pose a security risk in departure and arrival terminals of airports and train stations or other public places where people gather, such as shopping malls. Automated detection of abandoned objects helps security personnel respond promptly and mitigate potential threats [105].

5) CROWD MONITORING

Edge video analytics can analyze crowd behavior in real-time, detecting crowd density, movement patterns, or abnormal behavior within a crowd. This enables security personnel to identify potential crowd-related incidents, such as overcrowding, stampedes, or aggressive behavior, and take appropriate actions to maintain public safety [9]. An expert system based on video surveillance has been proposed in [106] with the aim of ensuring efficient monitoring of shopping malls by detecting abnormal behavior in real-time. The police can be alerted quickly when a shooting occurs or a dangerous person with a criminal history appears. Also, the authors in [107] addressed the challenges of accurate monitoring and counting of crowds in public spaces or events. Their practical solution leverages computer vision algorithms to analyze video data and extract meaningful information about the dynamics of a crowd.
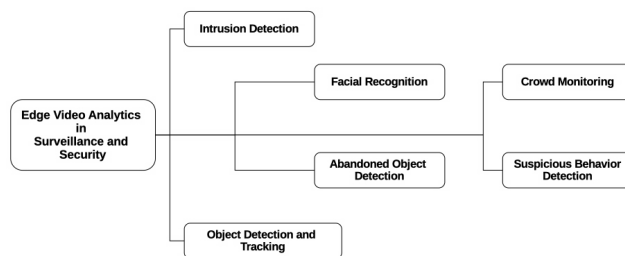

FIGURE 8. Use cases of edge video analytics in surveillance and security.

6) SUSPICIOUS BEHAVIOR DETECTION

Advanced behavior recognition algorithms can detect suspicious activities or behavior patterns by analyzing video feeds at the edge. This includes loitering, unauthorized access attempts, or suspicious gestures. Real-time alerts can be generated, enabling security personnel to respond swiftly and prevent security incidents. The authors in [108] addressed the need for efficient and accurate anomaly detection in surveillance systems, which play a crucial role in identifying abnormal events or behaviors. Their proposed solution for real-time anomaly detection integrates CNN features from video frames into bi-directional LSTM networks. It outperforms traditional methods by achieving remarkable accuracy.

Numerous research initiatives within the European Union (EU) have effectively combined video analytics and security, exemplifying a dynamic synergy [105]. Noteworthy among these is the APPS (Advancing Plug & Play Smart Surveillance) project, a concerted effort aimed at facilitating the seamless implementation of plug-and-play solutions, heightening the efficacy of intelligent decision-making, and engendering resilient communication mechanisms. In a parallel vein, the EWISA (Early Warning for Increased Situational Awareness) project diligently endeavors to construct an advanced early warning system, fostering a heightened level of situational awareness. Moreover, the INSIST (Integrated service delivery for Citizens' Safety and Comfort) project endeavors to create a smart environment wherein video surveillance and lighting management converge, fostering both public safety and a sense of comfort.

Since video data extracted from cameras in public places may contain sensitive data such as people's faces or car license plates, users should comply with security and privacy laws to minimize the potential risks. In [109], Lachner et al. identified and evaluated the impact of factors that can be adjusted in AI-assisted privacy for video analytics based on a face-blurring pipeline and edge computing. In [110], Grambow et al. explored the geographic distribution of fog computing servers to ensure the privacy and confidentiality of citizens' data. In addition, the work in [111] seeks to preserve privacy and security while accessing the information by implementing a privacy-preserving edge computing system for AI-assisted video analytics that restricts the app's access to a limited subset of the video stream data.

B. RETAIL AND MARKETING

Edge video analytics for retail and marketing is an active area of research. Many efforts focus on using computer vision and machine learning techniques to analyze retail video data. These studies primarily focus on using video data to improve sales and customer experience and gain insights into customer behavior. Video analytics in retail stores represents a widespread utilization wherein the tracking of customers and products is conducted. Such an approach facilitates the comprehensive comprehension of customer-product interactions, popular merchandise identification, and identification of the store's most frequented areas. The resultant insights hold significant potential for enhancing store layout, product positioning, and personnel allocation. Another area of research leverages video analytics to improve the customer experience. This includes using video data to monitor queues, identify customer needs and ensure staff provides excellent service.

Video analytics in marketing and advertising reveals customer demographics and interactions with digital signage. This knowledge fuels effective campaigns and targeted approaches, improving sales and customer experience. It also provides insights into retail and marketing behavior. Challenges include privacy concerns and lighting variations, but ongoing research aims to overcome them.

The supply chain and retail sectors are rapidly evolving, investing in research and technology for improved performance. Retailers prioritize customer satisfaction and data-driven decisions through e-commerce, loyalty programs, and optimized logistics. Autonomous decision-making predicated on real-time data analysis is poised to revolutionize the landscape of smart factories and logistics. Edge video analytics, driven by data sovereignty and reliability, supports multiple departments by tracking products in logistics, retail, and marketing. In the realm of AI, diverse models can be deployed on shared devices and nodes to tackle distinct tasks. Zhou et al. in [112] proposed a solution wherein complex edge AI models are partitioned into subtasks, enabling the simultaneous processing of shared data. Although this strategy may give rise to intricate algorithms, it yields superior calibration of investments and facilitates more precise measurements of marketing campaigns, thus enhancing their overall effectiveness.

According to Sandeep, [113], blockchain is widely used within the retail supply chain; it has several advantages in processing data generated by sensors. It can make shipping, tracking, and invoicing efficient. The convergence of blockchain and edge video analytics is being further explored by researchers, such as Jahid et al. [114]. Retail blockchain encompasses product tracking from the supplier or manufacturing unit, through the warehouse and retailer, to delivery to the end consumer. The entire process is fully automated, and the data is secured, helping develop new methods to improve connectivity mechanisms, introduce new services, and gain more insights.

The future of retail is product personalization, and customers need anticipation. With human understanding critical to future business direction, dynamic dashboards like those seen in some futuristic movies are not that far off. Video analytics in retail can spawn numerous applications to understand customer preferences and analyze foot traffic to create new sales opportunities. Video analytics is also being integrated into Industry 4.0 when it comes to detecting the lifecycle of equipment in technical areas, lean management, and maintenance schedules. The potential of IoT devices for last-mile optimization is enormous.

Cutting-edge research focuses on real-time processing to engage customers before leaving the store. The 2020 St. John University business review [115] highlighted case studies utilizing advanced 3D sensors to identify each customer and their unique profile. Real-time processing leverages past preferences and bill analysis for personalized product recommendations. With customer consent, immediate shelf replenishment occurs via internal data warehouse communications. In typical retail settings, linked sensors enable edge video analytics and integrated data analysis for proactive customer behavior anticipation. Alerts aid in identifying marketing successes or failures. However, processing the vast amount of data remains challenging, prompting research on collaborative edge cloud computing and addressing key IoT challenges. Also, privacy and ethical concerns arise when utilizing social media profiles and video surveillance data.

C. TRAFFIC AND TRANSPORTATION

In the context of traffic and transportation, edge video analytics can provide several benefits and applications. As depicted in Figure 9, here are the typical use cases in this context:

1) TRAFFIC MONITORING AND MANAGEMENT

Video analytics has proven to help traffic management in smart cities tremendously [116]. Insufficient traffic management measures in response to abrupt surges in highway and urban traffic can result in increased risks of accidents and congestion. Hence, edge video analytics solutions can be vital in this scenario. Indeed, real-time traffic analysis can help adjust traffic signal control systems dynamically and monitor congestion on roads and highways. It can also detect dangerous situations such as a vehicle moving erratically, driving in the opposite direction, or stopping at an unauthorized place. In the event of an accident, these systems make it possible to collect evidence during a trial. Incorporating edge analytics capabilities, camera systems exhibit the capacity to detect and track vehicles, record their traffic volume in specific regions, and examine the dynamics of traffic flow. Counting vehicles or distinguishing between various types, such as cars, trucks, buses, and cabs, provides high-quality statistics that shed light on traffic [34]. Automatic license plate recognition identifies cars committing an infraction or detects a stolen vehicle or a vehicle used for a crime, thanks to real-time search. The

acquired information proves invaluable in fine-tuning the synchronization of traffic signals, detecting bottlenecks, and making informed decisions with regard to traffic management strategies. Traffic engineers can then spot the peak hours of traffic congestion, closely examine traffic dynamics at intersections, and efficiently devise evidence-based strategies to expand or improve the roadway network.

Also, there are many use cases of edge video analytics to provide better public transportation services, also known as smart mobility. For instance, the routing of buses and trains should rely on actual demand for transportation, not just predefined routes with a fixed frequency. L. Cui et al. [117] studied the challenges of uploading video content by thousands of vehicles in terms of bandwidth consumption. They proposed an algorithm and queuing model for bus stops and used data from Shenzhen City in China to test the model. Their solution can be used to monitor vehicles and pedestrians separately, analyze bus stops and their relevance, and collect statistics at peak times to improve traffic management. Other valuable data can be collected via sensors and IoT devices, enabling more advanced services.

In their work, Albreem et al. [118] introduced an innovative edge visual sensor that effectively gathers data from the Agnosticity framework. This data, in turn, undergoes onboard computation facilitated by the real-time tracking algorithm known as SORT [119]. Singh et al. [120] have thoroughly investigated Internet of Things (IoT) technologies. They focused particularly on IoT's profound impact in the area of sustainable rail transportation Their study delved into using the Message Queuing Telemetry Transport (MQTT) protocol for seamless equipment communications, ultimately fostering autonomous decision-making capabilities. These transformative applications are continuously evolving and eagerly spreading through various transportation domains that extend beyond the confines of the rail sector to include dynamic sectors such as shipping.

2) INCIDENT DETECTION
By analyzing video streams at the edge, video analytics algorithms can detect various traffic incidents such as accidents, breakdowns, or pedestrians crossing in unauthorized areas [121], [122], [123]. When an incident is detected, alerts can be generated immediately, enabling prompt response from traffic authorities or emergency services. This helps in reducing response time and improving overall safety.

3) TRAFFIC VIOLATION DETECTION
Edge video analytics extends its capabilities to encompass traffic enforcement, detecting violations such as red light infractions, parking violations, and speeding. By analyzing video footage in real time, the system can automatically detect these violations and generate alerts or notifications for enforcement agencies [27]. This helps in enhancing traffic law enforcement and promoting safer driving behaviors. The authors in [124] addressed the need for efficient and

real-time recognition of anomalous vehicles, such as stolen or suspicious vehicles, in urban traffic. They proposed a solution that leverages edge video analytics to perform the recognition tasks closer to the data source, reducing latency and improving system performance.

4) PARKING MANAGEMENT
Edge video analytics can assist in parking management by monitoring parking spaces in real time. Cameras equipped with edge analytics can identify available parking spots, detect unauthorized parking, and provide occupancy information. Collected data can be used to guide drivers to available parking spaces, optimize parking space utilization, and prevent illegal parking. For example, the authors in [125] addressed the challenges of roadside occupation management, such as illegal parking or unauthorized use of dedicated spaces. They proposed a solution that employs computer vision algorithms to analyze video data captured from roadside cameras. The system utilizes object detection and tracking algorithms to identify and track vehicles and other objects occupying roadside spaces. It can detect and categorize various occupation scenarios, such as illegal parking, loading/unloading, or unauthorized usage.

5) INTELLIGENT TRANSPORTATION SYSTEMS (ITS)
Edge video analytics can be integrated into larger Intelligent Transportation Systems. By processing video data at the edge, traffic management systems can have near real-time access to valuable insights and actionable information. This enables adaptive traffic control systems, dynamic route guidance, and improved overall transportation efficiency. X. Zhou et al. [126] focused on data abstraction strategies and defined intelligent transportation systems (ITS) into five main components: traffic sensing, congestion management, data monitoring, communication, and control. Usually located at fixed sites, cameras can be assigned to specific tasks based on location and interest. During the sixth edition of the AI City Challenge [127], worldwide researchers gathered to optimize and implement algorithms. The detection-tracking-counting method remains one of the most powerful and cost-effective techniques.

In transportation, efforts are underway to improve computer technologies for less human-supervised and more autonomous decision-making. Autonomous vehicles are one of the most evolving concepts using edge video analytics. Vehicle equipment includes adapted sensors and devices for video capture. Huh et al. noted in [128] that Light Detection and Ranging (LIDAR) is currently the most advanced sensor that generates gigabytes of data per second. A camera as a single device can generate an enormous amount of information about movements, objects, interactions, and weather conditions. Compared to fixed cameras, cameras in vehicles pose a greater challenge for video analytics due to their motion. Javaid et al. [129] propose an edge analytics conversion of video frames into a discrete Markov decision
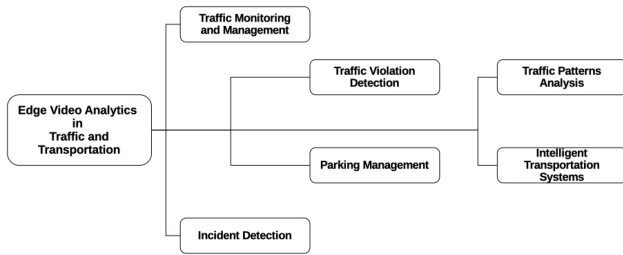
FIGURE 9. Use cases of edge video analytics in traffic management and transport.

process. Since the vehicle speed is constantly changing, the calibration parameters need to be further developed. Al-Ansi et al. estimated in [130] that a single autonomous vehicle could generate 30 terabytes of data daily. Deep reinforcement learning contributes to dynamic object detection based on the appearance of changing road types, environments, bicycles, people, pets, etc., to improve autonomous driving decisions.

It is fair to point out that many challenges still lie ahead for intelligent transportation technologies. Ke et al. [131] described several near-crash detections while designing algorithms to process and analyze dashcam video data. The algorithms should adapt to blurry video due to weather conditions or other causes and filter out irrelevant motion in dynamic environments. The development of smart IoT devices necessitates innovations in wireless communications and sensor technologies. Notably, The integration of such innovations becomes indispensable in creating intelligent devices, such as smart cameras, which play a pivotal role in deciphering the probable conduct of road users [132]. Leveraging advanced filtering techniques, these smart cameras emerge as indispensable tools in unraveling and comprehending the intricate behavioral patterns exhibited by individuals navigating the roadways.

Future research directions of ITS include edge cloud through B5G/6G [133], which promises extremely low latency and high bandwidth. To this end, researchers are also looking at multi-agent reinforcement learning (RL) [134] to explore complex traffic situations. The energy transition and the introduction of environmentally friendly technologies are being considered in computing devices. Using renewable energy sources to power IoT devices has already been studied by researchers such as M. Albreem et al. [118], who named it the Green Internet of Things (GIoT). This critical aspect needs to be further developed.

D. INDUSTRIAL AND MANUFACTURING SECTORS

In industrial environments, the fusion of computer vision and machine learning unlocks a realm of possibilities, enabling comprehensive analysis of video data. These dynamic technologies empower organizations to optimize operational efficiency, elevate quality control standards, and strengthen safety measures within their domains. Video analytics

finds common applications in monitoring industrial equipment performance, including robots and conveyors. This enables prompt identification of issues like jams or malfunctions, facilitating overall performance enhancement. Another research area focuses on leveraging video analytics for manufacturing quality control. Video data detect defects during production and ensures a smooth and efficient manufacturing process.

Several research efforts heavily delve into video analytics for safety monitoring in industrial and manufacturing environments, encompassing hazard detection, improper protective equipment, and identification of unsafe behavior. These efforts demonstrate the potential of video analytics to enhance efficiency, quality control, and safety. However, many challenges still require careful attention and a solution. They include poor lighting conditions, occlusions, and lighting variations.

The proliferation of IoT in the industrial sector permits monitoring equipment and collecting massive amounts of data that needs to be harnessed. Many research initiatives focused on optimizing storage resources and computing latency. Jianyu et al. [139] designed a real-time monitoring data system prototype integrating Network Functions Virtualization (NFV) and Software-Defined Networking (SDN). They want to prove the great potential of IoT experimentally in large-scale urban applications. Yi-Yun et al. [140] studied the concept of Factory Of the Future (FoF). They used real-time video analytics and developed a cloud-edge computing architecture with 5G wireless technology to deal with the resource allocation problem.

The industry is currently witnessing ongoing advancements in the domain of Video Analytics applications, offering diverse avenues for research and exploration. Among the prominently investigated areas is the notion of anomaly detection, which holds immense potential. Significant technical progress has been made in the realm of detecting precise anomalies utilizing video surveillance cameras. Devashree [141] investigated this concept and stated that there are many methodologies for anomaly detection. There is a growing trend to explore time-critical anomaly detection using Edge devices. Nevertheless, many challenges still require careful attention. They include training model improvement and hardware and software optimization.

E. HEALTHCARE

Edge video analytics has the potential to play a crucial role in healthcare by leveraging advanced technologies to analyze video data in healthcare facilities closer to the source of data generation. It involves using intelligent algorithms and machine learning models to process and interpret video streams captured from surveillance cameras, wearable devices, or other video-enabled devices in healthcare settings. Several studies show the important role of video analytics in healthcare. Here are some of the applications of edge video analytics in healthcare (see Figure 10).

TABLE 8. Summary of research studies on traffic and transportation based on video analytics and IoT.

Reference	Year	Study target and addressed issues	Methodology and algorithms
[120]	2022	• Holistic survey of IoT technologies in railway operations • IoT applications in maglev trains • Green IoT applications	• Fog computing • Radio Frequency Identification System (RFID) technology
[135]	2022	• Overview of emerging Technologies for Smart Cities Transportation • Geoinformation and communication technology (GeoICT) • Recent trends using integrated deep learning	• Big Data Architectures review • Data-driven transportation systems • Transportation Empowered by AI
[124]	2022	• Urban transportation monitoring • Defining database of normal vehicle activity patterns on the road • Identifying anomalous vehicle behaviors in real time	• SurMon: Smart Urban traffic Monitoring scheme employing drones • multidimensional Singular Spectrum Analysis (mSSA) technique
[136]	2022	• Processing and analyzing the data at the edge of the network • Monitoring parameters of the public traffic management system	• Empirically intelligent XGboost (EIXGB)-enabled logistic transportation system • Machine learning techniques
[137]	2022	• Vehicle Counting in Intelligent Transportation System (ITS) • Counting performance improvement by exploiting videos captured from multiple city cameras	• CityCam-to-Edge cooperative learning framework • Deep learning architecture • Vehicle counting models, F2F-M and O2O-M
[131]	2021	• Real-time near-Crash car detection • Algorithm models development	• Edge computing with traffic video analytics • TTC and horizontal motion estimations • Experiments and testing on cars and buses
[126]	2021	• Intelligent Transportation Systems (ITS) • Traffic Flow Detection • Future Research Directions	• Data Pre-Processing • Traffic Pattern Analysis • ITS Sensing, LiDAR and Camera
[117]	2021	• Video uploading problem for moving buses in public transportation systems • Video uploading delay and cost • Experimental results	• Video uploading in dedicated access points (AP) for video uploading • Water filling placement (WFP) algorithm • Gradient descent (GD) based algorithm
[138]	2020	• Real-time urban traffic planning • Experimental application on traffic video stream data	• MEMTV, a multi-level edge computing model • Big data machine
[132]	2020	• Investigating on-camera filtering • Exploring Reducto, a system that adapts filtering decisions according to feature type, query accuracy, and video content	• Video analytics, deep neural networks, object detection • Experimentation of Reducto on a variety of videos
[116]	2019	• Case study of Liverpool Smart Pedestrians project • Real-Time Traffic Monitoring • Designing an edge-computing device	• Deep neural networks • SORT tracking algorithm

1) REAL-TIME MONITORING AND ALERTING

The utilization of edge video analytics enables the continual surveillance of video feeds derived from diverse sources, including patient rooms, operating rooms, and waiting areas, in order to discern and carefully examine critical events in real time. This enables healthcare professionals to promptly

respond to emergencies, monitor patient safety, and identify potential risks or abnormalities [142]. Also, to assist child-care providers and nurseries in monitoring and analyzing the behavior of babies in smart homes or health centers, an intelligent IoT system with image sensors has been proposed [143] based on the analysis of baby videos and control charts to detect the baby's movements and send alerts when abnormal behavior occurs. As described in [108], an intelligent deep feature-based anomaly detection framework based on video image series and a CNN model was presented to detect anomalous events in video surveillance and reduce human work. Recently, in [144], the authors proposed an autism early detection system based on machine learning algorithms and video surveillance of children at home to detect abnormal events in a real usage context.

2) FALL DETECTION AND PREVENTION

By analyzing video streams, edge video analytics can detect falls or unusual movements of patients in hospitals or nursing homes. Healthcare personnel can be notified quickly through the generation of immediate alerts, thereby facilitating the provision of timely assistance and the prevention of subsequent injuries. Falls are among the most important health problems, especially for elderly people living alone. Therefore, several works based on video surveillance analysis have been conducted to solve this problem and provide quick first aid in emergencies. In [145], the authors implemented a non-invasive system based on RNNs and video streams to monitor fall detection in elderly living alone. The authors in [146] conducted research on monitoring falls on furniture in elderly people living at home or in hospitals using deep learning, R-CNN, activity features, and video scene analysis. The authors in [147] developed a smart home IoT system based on video stream analysis and feedback Optical Flow CNN to detect motion and recognize gestures and fall events.

3) PATIENT SAFETY AND SECURITY

Edge video analytics can enhance patient safety and security by monitoring restricted areas, tracking unauthorized access, and identifying potentially dangerous situations. It can also detect aggressive behavior or unusual activities, ensuring a safe environment for patients and staff. For example, enhanced care for the elderly and patients with disabilities requires a fast and efficient service. Therefore, video surveillance systems must respond promptly by sending real-time alerts and notifications to healthcare givers. To this end, the authors in [148] proposed an IoT-based intelligent system called "Cloud-based Object Tracking and Behavior Identification System (COTBIS)" to meet the need for a robust and secure healthcare video surveillance system. They used IoT and edge computing to quickly and easily analyze real-time data, reduce bandwidth, and decrease response time between video camera devices and the cloud server. They also used Deep Convolutional Neural Networks for object recognition, detection, and case activity classification of patients.
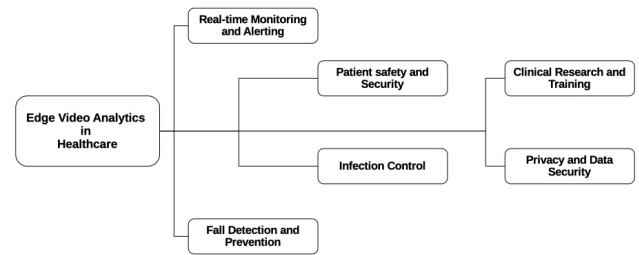

FIGURE 10. Use cases of edge video analytics in healthcare.

Various research efforts have used other types of video stream content analysis algorithms to improve intelligent video surveillance. These include computer vision algorithms, artificial neural networks (DNN, RNN, Deep Boltzmann Machine (DBM), and Deep Belief Networks (DBN), etc.) [149]. In [150], the authors used Region-based Convolutional Neural Networks (R-CNNs) to accelerate the detection of moving objects in a real-time video surveillance application. As reported in surveys ([151], [152]), Cloud Computing, Edge Computing, and Fog Computing can improve the robustness of the video surveillance system. In [153], the authors found that biometric features based on the time domain help optimize security in real-time smart health applications based on the Internet of Medical Things (IoMTs).

4) INFECTION CONTROL

Video analytics can aid in infection control efforts within healthcare facilities. Through the observation of hand hygiene, mask-wearing, and social distancing adherence, the system possesses the ability to identify areas of potential improvement and elevate compliance with infection control protocols. During the Coronavirus pandemic, an intelligent video early warning system was developed based on face recognition storage and data transmission between the edge device and the cloud server to ensure healthy and safe conditions and increase the safety of workers [154].

5) CLINICAL RESEARCH AND TRAINING

Edge video analytics can be utilized for clinical research, such as observing patient behavior or analyzing treatment outcomes [155], [156], [157], [158]. It can also support healthcare training by capturing video footage for educational purposes, allowing medical professionals to review and learn from real-life scenarios.

6) PRIVACY AND SECURITY

Edge video analytics focuses on processing and analyzing data locally, reducing the need for transmitting sensitive video streams to remote servers. This approach helps address privacy concerns and enhances data security since critical information can be processed and stored within the healthcare facility's infrastructure [111].
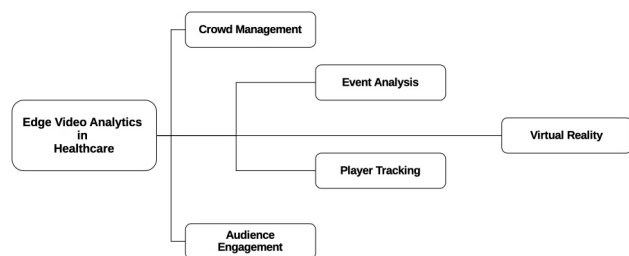
FIGURE 11. Use cases of edge video analytics in Sports and Entertainment.

F. SPORTS AND ENTERTAINMENT

The field of edge video analytics for sports and entertainment has witnessed a substantial amount of research in recent years. As shown in Figure 11, some examples of research works in this field include:

- *Crowd management.* Edge video analytics for crowd management aims to analyze video data from cameras at sports and entertainment venues to detect and track people, estimate crowd density, and identify potential hazards or security threats [159], [160].

- *Event analysis.* The objective of edge video analytics for event analysis is to develop algorithms and systems that are capable of scrutinizing video data emanating from sports and entertainment events with the aim of extracting significant details such as player tracking, ball tracking, and event highlights. This extracted information can subsequently be employed to enhance the fan experience and bolster revenue [161], [162], [163].

- *Player tracking.* Research on edge video analytics for player tracking aims to develop systems to track players and balls in real-time in sports such as basketball, soccer, and American football [164], [165], [166], [167]. These systems can provide insights on players' performance, the team's strategy, and overall match analysis.

G. EDUCATION

Numerous scholarly investigations identified within the scope of this review have delved into the implementation of edge video analytics within educational environments.

Zhou et al. [168] introduced an innovative video feature framework employing machine learning and computer vision techniques. This framework aims to forecast and comprehend online video consumption through a content-based lens. By applying the framework to distinct datasets, the authors validate its precision in predicting individual-level consumer behavior and overall video popularity across these diverse contexts. Furthermore, the authors elucidate the potential for their findings and methodologies to propel advancements in management and marketing research.

The authors in [169] use machine learning and virtual reality to analyze and improve teaching videos for oil painting art. They propose a deep learning-based object extraction fusion method and apply aesthetic criteria to filter out low-quality objects. The images are enhanced through saliency expansion, contour matching, and style migration. Virtual reality technology enhances art appreciation and student learning experiences, improving aesthetic quality.

In their paper, Li et al. [170] used a classroom-based video analysis research framework to conduct a micro-level empirical study on the relationship between teaching behavior, media application, and teacher knowledge structure in the smart classroom. They developed an analysis coding system to slice the classroom teaching video, collate the data, and analyze the statistics according to the teacher's teaching behavior, media use, and knowledge structure in the smart classroom. The study demonstrated that the integration of media within the intelligent classroom can have a favorable impact on student learning. However, the degree of this impact is contingent upon the media's utilization and its compatibility with the teacher's pedagogical approach and cognitive framework. Teachers with a strong knowledge structure are likelier to use media effectively in the smart classroom. The utilization of media can help them clearly understand the educational requirements of their students, thereby enabling them to offer more personalized guidance.

The authors in [171] developed Engage AI, a system to help teachers assess student engagement and attention during online teaching in the COVID-era. The system uses video-based machine learning models to detect emotions like happiness and neutrality, as well as drowsiness. It aggregates this data in a dashboard that instructors can view in real-time. This allows instructors to adjust their teaching to keep students engaged.

To enhance learning in computer laboratories, the authors in [172] propose a system that recognizes and localizes student actions in still images from CCTV videos. The method combines YOLOv3, a real-time object detection technology, with image template matching for efficient video processing. The authors create the STUDENT ACTION dataset using CCTV frames from a university computer laboratory to address the lack of a standard dataset. Their proposed method demonstrates excellent performance in recognizing various actions, particularly those with more samples.

The authors in [173] explored current classroom teaching interaction forms in technology-rich environments and identified their deficiencies. They analyzed primary math classrooms using interactive whiteboards, interactive televisions, or mobile terminals using the ITIAS (Information Technology-based Interaction Analysis System) coding system. They selected 20 teaching cases as samples and applied computer vision for video analysis. Four aspects were examined and concerned classroom atmosphere, teacher-student interaction, student-student interaction, and human-technology interaction. Cluster analysis revealed three interaction patterns: immediate interaction, waiting interaction, and shallow interaction.

In [174], the authors presented a visual analysis of clickstream data generated by learner interaction with course videos in MOOCs. The aim is to predict learner performance and enable instructors to make measures for timely
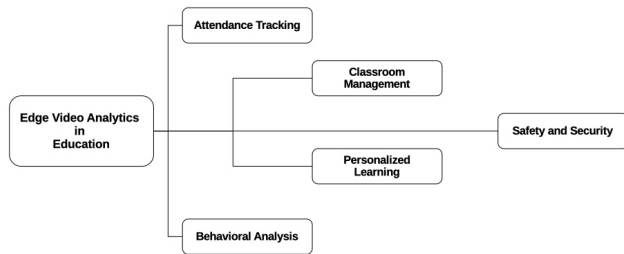
FIGURE 12. Use cases of edge video analytics in education.

intervention. The paper uses an LSTM network on implicit features extracted from video-clickstream data to predict learners' performance. The authors stated that their proposed LSTM model outperforms baseline Deep learning (GRU) and simple recurrent neural networks by an accuracy of 90.30% in the ''Mining of Massive Datasets'' course, and the ''Automata Theory'' accuracy is 89%.

Overall, edge video analytics can be used in an educational context in a variety of ways, as depicted in Figure 12.

VII. CHALLENGES AND OPEN ISSUES

This section addresses **RQ-4**. In the field of edge video analytics, several challenges and open issues still require careful attention. They are discussed in this section.

A. REAL-TIME PROCESSING

The imperative for real-time processing presents a significant obstacle within the realm of edge video analytics. The timely analysis and subsequent actions upon video data necessitate near-instantaneous processing, an endeavor fraught with challenges due to the sheer magnitude of data to be processed and the inherent constraints on processing power encountered in edge devices. Accommodating the performance requirements of video analytics applications becomes pivotal in order to fulfill the demands of specific use cases and optimize the user experience. The fulfillment of accuracy, latency, and throughput requisites assumes utmost significance for models employed in AI applications. Furthermore, developers of such models must be aware of the hardware constraints that may exert influence over the dimensions of model size and memory requirements. Striking the optimal equilibrium between model accuracy, inference speed, and model size often proves to be a formidable task for developers.

B. LIMITED RESOURCES

Edge devices, encompassing an array of IoT devices and cameras, are commonly characterized by their inherent constraints pertaining to processing power, memory, and storage capabilities. Such limitations render the execution of intricate algorithms and the storage of voluminous data on these devices a formidable undertaking. The development of algorithms specifically tailored to operate within the confines of resource-constrained devices assumes paramount significance for the proliferation of Edge AI applications, facilitating the utilization of these devices even in remote or challenging terrains that are otherwise difficult to access.

Most edge hardware has limited processing power, leading to inference constraints on object detection models. Efficient, low-memory, and low-power architectures are crucial for edge devices, considering heat dissipation limitations.

C. POWER AND ENERGY EFFICIENCY

Edge devices are often battery-powered or have limited power supplies, so power and energy efficiency are critical considerations. Algorithms and hardware must be designed to minimize power consumption while providing accurate and reliable results [175]. High accuracy and low power consumption are paramount concerns in the development of algorithms for video analytics applications such as object detection.

D. PRIVACY AND SECURITY

Edge video analytics encompasses a host of critical privacy and security considerations, primarily stemming from collecting and analyzing video data within sensitive environments, which frequently includes personal information. Safeguarding the privacy and security of such data emerges as a significant challenge, demanding meticulous attention and concerted efforts within this realm. Federated learning is a distributed approach to machine learning that addresses these concerns [176], [177], [178]. Federated learning necessitates the retention of data on edge devices as opposed to centralized storage in a singular location. The model is trained locally on those devices before the updated parameters are aggregated and shared among all participants. The reduction of sensitive data transmission and centralized storage in edge video analytics can serve as a solution to address privacy and security concerns. This approach can significantly mitigate the risk of security breaches and hacking attempts that may exploit the vulnerabilities of a centralized storage location.

E. SCALABILITY

A tough challenge awaits regarding the scalability of edge video analytics solutions. The exponential growth in the volume of data to process and store is directly attributable to the rapid proliferation of devices and streams. Developers seek scalable solutions without breaking the bank, opting to process multiple high-resolution video streams on a single device. This cost-effective approach defies convention but introduces an additional hurdle in achieving commercial viability at a large scale.

F. ROBUSTNESS

Edge video analytics systems must exhibit persistent resilience in various environmental conditions. They should be undeterred by challenging situations like low light, fog, and rain.

G. ADAPTABILITY

Edge video analytics systems must adapt to diverse scenarios involving different cameras, lighting conditions, and data types (video, audio, and sensor data).

H. HUMAN-IN-THE-LOOP

Human input and feedback are important for many edge video analytics applications, such as video surveillance, to improve the performance of the system and make it more accurate than with purely automated systems. In some cases, edge video analytics applications may require human input to train and fine-tune machine learning models since these models often rely on large amounts of labeled data. Humans can provide this labeled data by annotating or verifying the outputs of the models. In other cases, edge video analytics applications may rely on human feedback to improve their accuracy and reliability. For example, if a video analytics system monitors a public space for security purposes, it may flag potential security incidents for human review. The human reviewer can then provide feedback on the accuracy of these flags, allowing the system to learn and improve over time. By incorporating human input and feedback, edge video analytics applications can better address the challenges of data privacy and security while also improving accuracy and reliability.

The remarkable advances witnessed in computing paradigms and artificial intelligence have paved the way for an exciting realm of possibilities, prompting the identification of numerous vital research areas poised to propel the field of edge video analytics toward unprecedented horizons in the foreseeable future. Some of the most promising areas include:

I. EXPLAINABLE AI

The development of algorithms capable of furnishing lucid explanations for their outcomes assumes paramount significance within the purview of video analytics applications, particularly in domains such as security and surveillance, wherein the ability to decipher and comprehend the algorithmic outputs holds unparalleled importance.

J. MULTI-MODAL FUSION

Developing algorithms that fuse information from multiple modalities (video, audio, sensor data) is crucial for enhancing edge video analytics in applications like surveillance and event detection.

K. ADVERSARIAL ROBUSTNESS

Developing algorithms that can resist attacks from adversarial examples will be necessary for many edge video analytics applications, particularly in security and surveillance, for which detecting and responding to attacks is essential.

L. EDGE-CLOUD COLLABORATION

Developing algorithms for effective collaboration between edge devices and the cloud will be necessary for many video analytics applications. The synergistic partnership between the edge and the cloud will facilitate the utilization of enhanced algorithms and more significant amounts of data while preserving minimal latency and power consumption.

VIII. CONCLUSION

In recent years, the field of Edge AI has surfaced as a technology with great potential for transforming video analytics. This technology presents a novel approach characterized by its ability to facilitate real-time processing, minimize latency, optimize privacy and security, and accommodate resource-limited settings. This article has provided a systematic review of the scholarly works on Edge AI-assisted video analytics in smart cities. It has undertaken the task of classifying the diverse research endeavors within a taxonomy. Additionally, it has thoroughly scrutinized the numerous models of artificial intelligence, as well as privacy-preserving techniques employed in edge video analytics. Furthermore, this review has described various applications of edge video analytics across a multitude of domains like surveillance, transportation, retail, manufacturing, healthcare, and education.

Nevertheless, several challenges still need to be addressed to realize Edge AI's potential for video analytics fully. These include enhancing the accuracy and robustness of algorithms, reducing power consumption and increasing processing speed, and ensuring that Edge AI technology can be easily integrated into video analytics systems close to where data is generated.

Despite these challenges, edge video analytics holds a promising future. The revolutionary potential of Edge AI in analyzing video data and facilitating decision-making is substantial. Researchers and practitioners must stay updated on advancements and strive to develop innovative solutions for video analytics.

REFERENCES

[1] T. Bai, J. Wang, Y. Ren, and L. Hanzo, "Energy-efficient computation offloading for secure UAV-edge-computing systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6074–6087, Jun. 2019.

[2] M. Li, N. Cheng, J. Gao, Y. Wang, L. Zhao, and X. Shen, "Energy-efficient UAV-assisted mobile edge computing: Resource allocation and trajectory optimization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3424–3438, Mar. 2020.

[3] Y. Lu, J. Zhang, Y. Qi, S. Qi, Y. Zheng, Y. Liu, H. Song, and W. Wei, "Accelerating at the edge: A storage-elastic blockchain for latency-sensitive vehicular edge computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 11862–11876, Aug. 2022.

[4] Y. Qu, H. Dai, H. Wang, C. Dong, F. Wu, S. Guo, and Q. Wu, "Service provisioning for UAV-enabled mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3287–3305, Nov. 2021.

[5] H. Watanabe, R. Yasumori, T. Kondo, K. Kumakura, K. Maesako, L. Zhang, Y. Inagaki, and F. Teraoka, "ContMEC: An architecture of multi-access edge computing for offloading container-based mobile applications," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 3647–3653.

[6] Y. Sha, M. Li, H. Xu, S. Zhang, and T. Feng, "Smart city public safety intelligent early warning and detection," *Sci. Program.*, vol. 2022, pp. 1–11, Jun. 2022.

[7] E. Maltezos, L. Karagiannidis, A. Dadoukis, K. Petousakis, F. Misichroni, E. Ouzounoglou, L. Gounaridis, D. Gounaridis, C. Kouloumentas, and A. Amditis, "Public safety in smart cities under the edge computing concept," in *Proc. IEEE Int. Medit. Conf. Commun. Netw. (MeditCom)*, Sep. 2021, pp. 88–93.

[8] N. Papadakis, N. Koukoulas, I. Christakis, I. Stavrakas, and D. Kandris, "An IoT-based participatory antitheft system for public safety enhancement in smart cities," *Smart Cities*, vol. 4, no. 2, pp. 919–937, Jun. 2021.

[9] Q. Zhang, H. Sun, X. Wu, and H. Zhong, "Edge video analytics for public safety: A review," *Proc. IEEE*, vol. 107, no. 8, pp. 1675–1696, Aug. 2019.

[10] M. A. Khan, E. Baccour, Z. Chkirbene, A. Erbad, R. Hamila, M. Hamdi, and M. Gabbouj, "A survey on mobile edge computing for video streaming: Opportunities and challenges," *IEEE Access*, vol. 10, pp. 120514–120550, 2022.

[11] R. Xu, S. Razavi, and R. Zheng, "Deep learning-driven edge video analytics: A survey," 2022, *arXiv:2211.15751*.

[12] H. Zhong, J. Shi, and M. Visontai, "Detecting unusual activity in video," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2004, pp. 1–6.

[13] A. Choudhary, M. Pal, S. Banerjee, and S. Chaudhury, "Unusual activity analysis using video epitomes and pLSA," in *Proc. 6th Indian Conf. Comput. Vis., Graph. Image Process.*, Dec. 2008, pp. 390–397.

[14] L. Kratz and K. Nishino, "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1446–1453.

[15] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[16] P. G. Lopez, A. Montresor, D. Epema, A. Datta, T. Higashino, A. Iamnitchi, M. Barcellos, P. Felber, and E. Riviere, "Edge-centric computing: Vision and challenges," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, Oct. 2015.

[17] A. J. Ferrer, J. M. Marquès, and J. Jorba, "Towards the decentralised cloud: Survey on approaches and challenges for mobile, ad hoc, and edge computing," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 1–36, Jan. 2019.

[18] S. Voghoei, N. Hashemi Tonekaboni, J. G. Wallace, and H. R. Arabnia, "Deep learning at the edge," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2018, pp. 895–901.

[19] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the Internet of Things with edge computing," *IEEE Netw.*, vol. 32, no. 1, pp. 96–101, Jan. 2018.

[20] A. Munir, E. Blasch, J. Kwon, J. Kong, and A. Aved, "Artificial intelligence and data fusion at the edge," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 7, pp. 62–78, Jul. 2021.

[21] Z. Zhao, K. Wang, N. Ling, and G. Xing, "EdgeML: An AutoML framework for real-time deep learning on the edge," in *Proc. Int. Conf. Internet-Things Design Implement.*, New York, NY, USA, May 2021, pp. 133–144.

[22] S. Lin, Z. Zhou, Z. Zhang, X. Chen, and J. Zhang, *Edge Intelligence in the Making: Optimization, Deep Learning, and Applications*. Cham, Switzerland: Springer, Jun. 2022.

[23] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee, and B. Yin, "Smart factory of industry 4.0: Key technologies, application case, and challenges," *IEEE Access*, vol. 6, pp. 6505–6519, 2018.

[24] N. Anumbe, C. Saidy, and R. Harik, "A primer on the factories of the future," *Sensors*, vol. 22, no. 15, p. 5834, Aug. 2022.

[25] B. Krishnakumar, K. Kousalya, R. S. Mohana, E. K. Vellingiriraj, K. Maniprasanth, and E. Krishnakumar, "Detection of vehicle speeding violation using video processing techniques," in *Proc. Int. Conf. Comput. Commun. Inform. (ICCCI)*, Jan. 2022, pp. 1–14.

[26] B. Isong, O. Khutsoane, and N. Dladlu, "Real-time monitoring and detection of drinkdriving and vehicle over-speeding," *Int. J. Image, Graph. Signal Process.*, vol. 9, no. 11, pp. 1–9, Nov. 2017.

[27] S. U. Khan, N. Alam, S. U. Jan, and I. S. Koo, "IoT-enabled vehicle speed monitoring system," *Electronics*, vol. 11, no. 4, p. 614, Feb. 2022.

[28] M. A. Khan, S. F. Khan, T. A. T. B. BaOmar, and A. R. M. A. B. Omar, "Development & implementation of smart vehicle over speeding detector using IoT," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 4, no. 2, pp. 170–175, 2019.

[29] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of artificial intelligence and machine learning in smart cities," *Comput. Commun.*, vol. 154, pp. 313–323, Mar. 2020.

[30] S. Paiva, M. Ahad, G. Tripathi, N. Feroz, and G. Casalino, "Enabling technologies for urban smart mobility: Recent trends, opportunities and challenges," *Sensors*, vol. 21, no. 6, p. 2143, Mar. 2021.

[31] E. Badidi, "Edge AI and blockchain for smart sustainable cities: Promise and potential," *Sustainability*, vol. 14, no. 13, p. 7609, Jun. 2022.

[32] V. K. Rathi, N. K. Rajput, S. Mishra, B. A. Grover, P. Tiwari, A. K. Jaiswal, and M. S. Hossain, "An edge AI-enabled IoT healthcare monitoring system for smart cities," *Comput. Electr. Eng.*, vol. 96, Dec. 2021, Art. no. 107524.

[33] M. M. Kamruzzaman, I. Alrashdi, and A. Alqazzaz, "New opportunities, challenges, and applications of edge-AI for connected healthcare in Internet of Medical Things for smart cities," *J. Healthc. Eng.*, vol. 2022, Feb. 2022, Art. no. 2950699.

[34] D.-L. Dinh, H.-N. Nguyen, H.-T. Thai, and K.-H. Le, "Towards AI-based traffic counting system with edge computing," *J. Adv. Transp.*, vol. 2021, pp. 1–15, Jun. 2021.

[35] Y. Liu, C. Yang, L. Jiang, S. Xie, and Y. Zhang, "Intelligent edge computing for IoT-based energy management in smart cities," *IEEE Netw.*, vol. 33, no. 2, pp. 111–117, Mar. 2019.

[36] W. Qi, Q. Li, Q. Song, L. Guo, and A. Jamalipour, "Extensive edge intelligence for future vehicular networks in 6G," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 128–135, Aug. 2021.

[37] E. Li, Z. Zhou, and X. Chen, "Edge intelligence: On-demand deep learning model co-inference with device-edge synergy," in *Proc. Workshop Mobile Edge Commun.*, Aug. 2018, pp. 31–36.

[38] T. Rausch and S. Dustdar, "Edge intelligence: The convergence of humans, things, and AI," in *Proc. IEEE Int. Conf. Cloud Eng. (IC2E)*, Jun. 2019, pp. 86–96.

[39] K.-C. Hsu and H.-W. Tseng, "Accelerating applications using edge tensor processing units," in *Proc. Int. Conf. High Perform. Comput., Netw., Storage Anal.*, New York, NY, USA, Nov. 2021, pp. 1–16.

[40] *Edge TPU—Run Inference at the Edge | Google Cloud*, Google, Mountain View, CA, USA, Nov. 2021, Accessed: Jan. 5, 2022.

[41] *NVIDIA Jetson Nano For Edge AI Applications and Education*, Nvidia, Santa Clara, CA, USA, Jan. 2022. Accessed: Jan. 5, 2022.

[42] *NVIDIA Jetson TX2: High Performance AI at the Edge*, Nvidia, Santa Clara, CA, USA, Jan. 2022. Accessed: Jan. 5, 2022.

[43] *Intel Movidius Vision Processing Units (VPUs)*, Intel, Santa Clara, CA, USA, Jan. 2022. Accessed: Jan. 5, 2022.

[44] Qualcomm Developer Network. *Qualcomm Neural Processing SDK*. Accessed: Jun. 5, 2023. [Online]. Available: https://developer.qualcomm.com/software/qualcomm-neural-processing-sdk

[45] OpenCv. (Mar. 2019). *Open Computer Vision library*. [Online]. Available: https://opencv.org/about/

[46] M. J. Page, "PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews," *BMJ*, vol. 372, p. n160, Mar. 2021.

[47] K. S. H. Ong, D. Niyato, C. So-In, and T. Friedrichs, "Accelerating low-cost edge-based real-time video analytics using task scheduling," in *Proc. IEEE 7th World Forum Internet Things (WF-IoT)*, Jun. 2021, pp. 598–603.

[48] G. Constantinou, C. Shahabi, and S. H. Kim, "Placement of DNN models on mobile edge devices for effective video analysis," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2021, pp. 207–218.

[49] Y. Wu, L. Liu, and R. Kompella, "Parallel detection for efficient video analytics at the edge," in *Proc. IEEE 3rd Int. Conf. Cognit. Mach. Intell. (CogMI)*, Dec. 2021, pp. 01–10.

[50] X. Li, B. Cho, and Y. Xiao, "Balancing latency and accuracy on deep video analytics at the edge," in *Proc. IEEE 19th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2022, pp. 299–306.

[51] J. Guo, S. Xia, and C. Peng, "VPPlus: Exploring the potentials of video processing for live video analytics at the edge," in *Proc. IEEE/ACM 30th Int. Symp. Quality Service (IWQoS)*, Jun. 2022, pp. 1–11.

[52] W.-J. Kim and C.-H. Youn, "A low-cost video analytics system with velocity based configuration adaptation in edge computing," in *Advances in Artificial Intelligence and Applied Cognitive Computing*, Cham, Switzerland: Springer, 2021, pp. 667–677.

[53] J. King and Y. C. Lee, "Distributed edge-based video analytics on the move," 2022, *arXiv:2206.14414*.

[54] Z. Xiao, Z. Xia, H. Zheng, B. Y. Zhao, and J. Jiang, "Towards performance clarity of edge video analytics," in *Proc. IEEE/ACM Symp. Edge Comput. (SEC)*, Dec. 2021, pp. 148–164.

[55] J. Lin, P. Yang, W. Wu, N. Zhang, T. Han, and L. Yu, "Edge learning for low-latency video analytics: Query scheduling and resource allocation," in *Proc. IEEE 18th Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, Oct. 2021, pp. 252–259.

[56] J. Wang, Z. Feng, Z. Chen, S. George, M. Bala, P. Pillai, S.-W. Yang, and M. Satyanarayanan, "Bandwidth-efficient live video analytics for drones via edge computing," in *Proc. IEEE/ACM Symp. Edge Comput. (SEC)*, Oct. 2018, pp. 159–173.

[57] M. N. Sadat, E. Vargas-Alfonso, and R. Dai, "Quality-aware video analytics in edge computing environments," in *Proc. IEEE 19th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2022, pp. 449–452.

[58] K. Chen, Y. Zhu, Z. Han, and X. Wang, "Adaptive cross-camera video analytics at the edge," in *Proc. IEEE 19th Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, Oct. 2022, pp. 394–402.

[59] L. Fick, S. Skrzyniarz, M. Parikh, M. B. Henry, and D. Fick, "Analog matrix processor for edge AI real-time video analytics," in *Proc. IEEE Int. Solid- State Circuits Conf. (ISSCC)*, vol. 65, Feb. 2022, pp. 260–262.

[60] Z. Hu, N. Ye, C. Phillips, T. Capes, and I. Mohomed, "mmFilter: Language-Guided video analytics at the edge," in *Proc. 21st Int. Middleware Conf. Ind. Track*, New York, NY, USA, Dec. 2020, pp. 1–7.

[61] S. P. Rachuri, F. Bronzino, and S. Jain, "Decentralized modular architecture for live video analytics at the edge," in *Proc. 3rd ACM Workshop Hot Topics Video Analytics Intell. Edges*, New York, NY, USA, Oct. 2021, pp. 13–18.

[62] D. V. Nguyen and J. Choi, "Toward scalable video analytics using compressed-domain features at the edge," *Appl. Sci.*, vol. 10, no. 18, p. 6391, Sep. 2020.

[63] X. Zeng, B. Fang, H. Shen, and M. Zhang, "Distream: Scaling live video analytics with workload-adaptive distributed edge intelligence," in *Proc. 18th Conf. Embedded Networked Sensor Syst.*, New York, NY, USA, Nov. 2020, pp. 409–421.

[64] B. Qian, Z. Wen, J. Tang, Y. Yuan, A. Y. Zomaya, and R. Ranjan, "OsmoticGate: Adaptive edge-based real-time video analytics for the Internet of Things," *IEEE Trans. Comput.*, vol. 72, no. 4, pp. 1178–1193, Apr. 2023.

[65] P. Yang, F. Lyu, W. Wu, N. Zhang, L. Yu, and X. S. Shen, "Edge coordinated query configuration for low-latency and accurate video analytics," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4855–4864, Jul. 2020.

[66] X. Dai, P. Yang, X. Zhang, Z. Dai, and L. Yu, "RESPIRE: Reducing spatial–temporal redundancy for efficient edge-based industrial video analytics," *IEEE Trans. Ind. Informat.*, vol. 18, no. 12, pp. 9324–9334, Dec. 2022.

[67] S. Zhang, C. Wang, Y. Jin, J. Wu, Z. Qian, M. Xiao, and S. Lu, "Adaptive configuration selection and bandwidth allocation for edge-based video analytics," *IEEE/ACM Trans. Netw.*, vol. 30, no. 1, pp. 285–298, Feb. 2022.

[68] J. A. Peris and V. Fodor, "Modelling multi-cell edge video analytics," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 1665–1671.

[69] T. Stone, N. Stone, P. Jain, Y. Jiang, K.-H. Kim, and S. Nelakuditi, "Towards scalable video analytics at the edge," in *Proc. 16th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2019, pp. 1–9.

[70] M. Ali, A. Anjum, M. U. Yaseen, A. R. Zamani, D. Balouek-Thomert, O. Rana, and M. Parashar, "Edge enhanced deep learning system for large-scale video stream analytics," in *Proc. IEEE 2nd Int. Conf. Fog Edge Comput. (ICFEC)*, May 2018, pp. 1–10.

[71] L. Zhang, G. Gao, and H. Zhang, "Towards data-efficient continuous learning for edge video analytics via smart caching," in *Proc. 20th ACM Conf. Embedded Networked Sensor Syst.*, New York, NY, USA, Jan. 2023, pp. 1136–1140.

[72] Y. Lu, S. Jiang, T. Cao, and Y. Shu, "Turbo: Opportunistic enhancement for edge video analytics," 2022, *arXiv:2207.00172*.

[73] Z. Wang, X. He, Z. Zhang, Y. Zhang, Z. Cao, W. Cheng, W. Wang, and Y. Cui, "Edge-assisted real-time video analytics with spatial–temporal redundancy suppression," *IEEE Internet Things J.*, vol. 10, no. 7, pp. 6324–6335, Apr. 2023.

[74] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[75] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[76] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 386–397, Feb. 2020.

[77] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[78] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Visiom—ECCV 2016*. Cham, Switzerland: Springer, 2016, pp. 21–37.

[79] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[80] X. Ran, H. Chen, X. Zhu, Z. Liu, and J. Chen, "DeepDecision: A mobile deep learning framework for edge video analytics," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2018, pp. 1421–1429.

[81] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010.

[82] P. Dubal, R. Mahadev, S. Kothawade, K. Dargan, and R. Iyer, "Deployment of customized deep learning based video analytics on surveillance cameras," 2018, *arXiv:1805.10604*.

[83] T. Tan and G. Cao, "Deep learning video analytics through edge computing and neural processing units on mobile devices," *IEEE Trans. Mobile Comput.*, vol. 22, no. 3, pp. 1433–1448, Mar. 2023.

[84] Y. Dong, G. Gao, R. Wang, and Z. Yan, "Collaborative video analytics on distributed edges with multiagent deep reinforcement learning," 2022, *arXiv:2211.03102*.

[85] S. Wang, C. Zhang, Y. Shu, and Y. Liu, "Live video analytics with FPGA-based smart cameras," in *Proc. Workshop Hot Topics Video Anal. Intell. Edges*, New York, NY, USA, Oct. 2019, pp. 9–14.

[86] P. Liu, B. Qi, and S. Banerjee, "EdgeEye: An edge service framework for real-time intelligent video analytics," in *Proc. 1st Int. Workshop Edge Syst., Anal. Netw.*, New York, NY, USA, Jun. 2018, pp. 1–6.

[87] G. Gad, E. Gad, K. Cengiz, Z. Fadlullah, and B. Mokhtar, "Deep learning-based context-aware video content analysis on IoT devices," *Electronics*, vol. 11, no. 11, p. 1785, Jun. 2022.

[88] J. Shao, X. Zhang, and J. Zhang, "Task-oriented communication for edge video analytics," 2022, *arXiv:2211.14049*.

[89] A. Galanopoulos, J. A. Ayala-Romero, D. J. Leith, and G. Iosifidis, "AutoML for video analytics with edge computing," in *Proc. IEEE Conf. Comput. Commun.*, May 2021, pp. 1–10.

[90] H. Barahouei Pasandi and T. Nadeem, "CONVINCE: Collaborative cross-camera video analytics at the edge," 2020, *arXiv:2002.03797*.

[91] C. Hu, R. Lu, and D. Wang, "FEVA: A federated video analytics architecture for networked smart cameras," *IEEE Netw.*, vol. 35, no. 6, pp. 163–170, Nov. 2021.

[92] A. Ben Sada, M. A. Bouras, J. Ma, H. Runhe, and H. Ning, "A distributed video analytics architecture based on edge-computing and federated learning," in *Proc. IEEE Int. Conf. Dependable, Autonomic Secure Comput., Int. Conf. Pervasive Intell. Comput., Int. Conf. Cloud Big Data Comput., Int. Conf. Cyber Sci. Technol. Congr.*, Aug. 2019, pp. 215–220.

[93] Y. Deng, T. Han, and N. Ansari, "FedVision: Federated video analytics with edge computing," *IEEE Open J. Comput. Soc.*, vol. 1, pp. 62–72, 2020.

[94] M. Mendieta, C. Neff, D. Lingerfelt, C. Beam, A. George, S. Rogers, A. Ravindran, and H. Tabkhi, "A novel application/infrastructure co-design approach for real-time edge video analytics," in *Proc. Southeast-Con*, Apr. 2019, pp. 1–7.

[95] B. Zhang, S. Shi, D. Wang, and C. Hu, "EPC: A video analytics system with efficient edge-side privacy control," in *Proc. 17th ACM Workshop Mobility Evolving Internet Archit.*, New York, NY, USA, Oct. 2022, pp. 37–42.

[96] F. Cangialosi, N. Agarwal, V. Arun, S. Narayana, A. Sarwate, and R. Netravali, "Privid: Practical, privacy-preserving video analytics queries," in *Proc. 19th USENIX Symp. Networked Syst. Design Implement.*, 2022, pp. 209–228.

[97] S. Zhang, Y. Feng, L. Bauer, L. F. Cranor, A. Das, and N. Sadeh, "'Did you know this camera tracks your mood?': Understanding privacy expectations and preferences in the age of video analytics," *Proc. Privacy Enhancing Technol.*, vol. 2021, pp. 282–304, Apr. 2021.

[98] J. L. Lavanchy, A. Vardazaryan, P. Mascagni, A. Consortium, D. Mutter, and N. Padoy, "Preserving privacy in surgical video analysis using artificial intelligence: A deep learning classifier to identify out-of-body scenes in endoscopic videos," 2023, *arXiv:2301.07053*.

[99] G. Giorgi, W. Abbasi, and A. Saracino, "Privacy-preserving analysis for remote video anomaly detection in real life environments," *J. Wireless Mobile Netw. Ubiquitous*, vol. 13, no. 1, pp. 112–136, 2022.

[100] W. Du, A. Li, P. Zhou, B. Niu, and D. Wu, "PrivacyEye: A privacy-preserving and computationally efficient deep learning-based mobile video analytics system," *IEEE Trans. Mobile Comput.*, vol. 21, no. 9, pp. 3263–3279, Sep. 2022.

[101] A. Gupta and P. Prabhat, "Towards a resource efficient and privacy-preserving framework for campus-wide video analytics-based applications," *Complex Intell. Syst.*, vol. 9, no. 1, pp. 161–176, Feb. 2023.

[102] A. Zaman, X. Liu, and Z. Zhang, "Video analytics for railroad safety research: An artificial intelligence approach," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 10, pp. 269–277, Dec. 2018.

[103] Q. Zhang, Q. Zhang, W. Shi, and H. Zhong, "Distributed collaborative execution on the edges and its application to AMBER alerts," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3580–3593, Oct. 2018.

[104] Y.-X. Liu, Y. Yang, A. Shi, P. Jigang, and L. Haowei, "Intelligent monitoring of indoor surveillance video based on deep learning," in *Proc. 21st Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2019, pp. 648–653.

[105] J. H. C. van Rest, "Surveillance use cases: Focus on video analytics," Joint Res. Centre, Eur. Commission's, Brussels, Belgium, Tech. Rep. JRC100401, 2015.

[106] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza, and J. Almazán, "Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7991–8005, Nov. 2015.

[107] K. H. Cheong, S. Poeschmann, J. W. Lai, J. M. Koh, U. R. Acharya, S. C. M. Yu, and K. J. W. Tang, "Practical automated video analytics for crowd monitoring and counting," *IEEE Access*, vol. 7, pp. 183252–183261, 2019.

[108] W. Ullah, A. Ullah, I. U. Haq, K. Muhammad, M. Sajjad, and S. W. Baik, "CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks," *Multimedia Tools Appl.*, vol. 80, no. 11, pp. 16679–16695, May 2021.

[109] C. Lachner, Z. Á. Mann, and S. Dustdar, "Towards understanding the adaptation space of AI-assisted data protection for video analytics at the edge," in *Proc. IEEE 41st Int. Conf. Distrib. Comput. Syst. Workshops (ICDCSW)*, Jul. 2021, pp. 7–12.

[110] M. Grambow, J. Hasenburg, and D. Bermbach, "Public video surveillance: Using the fog to increase privacy," in *Proc. 5th Workshop Middleware Appl. Internet Things*, New York, NY, USA, Dec. 2018, pp. 11–14.

[111] C. Lachner, T. Rausch, and S. Dustdar, "A privacy preserving system for AI-assisted video analytics," in *Proc. IEEE 5th Int. Conf. Fog Edge Comput. (ICFEC)*, May 2021, pp. 74–78.

[112] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019.

[113] S. Jagtap, F. Bader, G. Garcia-Garcia, H. Trollman, T. Fadiji, and K. Salonitis, "Food logistics 4.0: Opportunities and challenges," *Logistics*, vol. 5, no. 1, p. 2, Dec. 2020.

[114] A. Jahid, M. H. Alsharif, and T. J. Hall, "The convergence of blockchain, IoT and 6G: Potential, opportunities, challenges and research roadmap," 2021, *arXiv:2109.03184*.

[115] L. Costantino and M. Pompella, "Fintech in light of the 2020 emergency: Excess innovations to the facts–from securitization to tokenomics, and more," *Rev. Bus.*, vol. 41, no. 1, pp. 1–12, 2021.

[116] J. Barthélemy, N. Verstaevel, H. Forehead, and P. Perez, "Edge-computing video analytics for real-time traffic monitoring in a smart city," *Sensors*, vol. 19, no. 9, p. 2048, May 2019.

[117] L. Cui, D. Su, Y. Zhou, L. Zhang, Y. Wu, and S. Chen, "Edge learning for surveillance video uploading sharing in public transport systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2274–2285, Apr. 2021.

[118] M. A. Albreem, A. M. Sheikh, M. H. Alsharif, M. Jusoh, and M. N. M. Yasin, "Green Internet of Things (GIoT): Applications, practices, awareness, and challenges," *IEEE Access*, vol. 9, pp. 38833–38858, 2021.

[119] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3645–3649.

[120] P. Singh, Z. Elmi, V. Krishna Meriga, J. Pasha, and M. A. Dulebenets, "Internet of Things for sustainable railway transportation: Past, present, and future," *Cleaner Logistics Supply Chain*, vol. 4, Jul. 2022, Art. no. 100065.

[121] Y. Yang, S. Lee, and J. Lee, "Video analytics-based real-time intelligent crossing detection system (RICDS): Killer app for edge computing," *Future Gener. Comput. Syst.*, vol. 133, pp. 84–94, Aug. 2022.

[122] C. Neff, M. Mendieta, S. Mohan, M. Baharani, S. Rogers, and H. Tabkhi, "REVAMP2T: Real-time edge video analytics for multicamera privacy-aware pedestrian tracking," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2591–2602, Apr. 2020.

[123] K. R. Sri Preethaa and A. Sabari, "Intelligent video analysis for enhanced pedestrian detection by hybrid metaheuristic approach," *Soft Comput.*, vol. 24, no. 16, pp. 12303–12311, Aug. 2020.

[124] N. Chen and Y. Chen, "Anomalous vehicle recognition in smart urban traffic monitoring as an edge service," *Future Internet*, vol. 14, no. 2, p. 54, Feb. 2022.

[125] G. T. S. Ho, Y. P. Tsang, C. H. Wu, W. H. Wong, and K. L. Choy, "A computer vision-based roadside occupation surveillance system for intelligent transport in smart cities," *Sensors*, vol. 19, no. 8, p. 1796, Apr. 2019.

[126] X. Zhou, R. Ke, H. Yang, and C. Liu, "When intelligent transportation systems sensing meets edge computing: Vision and challenges," *Appl. Sci.*, vol. 11, no. 20, p. 9680, Oct. 2021.

[127] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, Y. Yao, L. Zheng, M. Shaiqur Rahman, A. Venkatachalapathy, A. Sharma, Q. Feng, V. Ablavsky, S. Sclaroff, P. Chakraborty, A. Li, S. Li, and R. Chellappa, "The 6th AI city challenge," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2022, pp. 3346–3355.

[128] J.-H. Huh and Y.-S. Seo, "Understanding edge computing: Engineering evolution with artificial intelligence," *IEEE Access*, vol. 7, pp. 164229–164245, 2019.

[129] M. Javaid, A. Haleem, R. Pratap Singh, S. Rab, and R. Suman, "Upgrading the manufacturing sector via applications of industrial Internet of Things (IIoT)," *Sensors Int.*, vol. 2, 2021, Art. no. 100129.

[130] A. Al-Ansi, A. M. Al-Ansi, A. Muthanna, I. A. Elgendy, and A. Koucheryavy, "Survey on intelligence edge computing in 6G: Characteristics, challenges, potential use cases, and market drivers," *Future Internet*, vol. 13, no. 5, p. 118, Apr. 2021.

[131] R. Ke, Z. Cui, Y. Chen, M. Zhu, H. Yang, and Y. Wang, "Edge computing for real-time near-crash detection for smart transportation applications," 2020, *arXiv:2008.00549*.

[132] Y. Li, A. Padmanabhan, P. Zhao, Y. Wang, G. H. Xu, and R. Netravali, "Reducto: On-camera filtering for resource-efficient real-time video analytics," in *Proc. Annu. Conf. ACM Special Interest Group Data Commun. Appl., Technol., Architectures, Protocols Comput. Commun.*, Jul. 2020, pp. 359–376.

[133] M. Ishtiaq, N. Saeed, and M. A. Khan, "Edge computing in IoT: A 6G perspective," 2021, *arXiv:2111.08943*.

[134] F. Wang, M. Zhang, X. Wang, X. Ma, and J. Liu, "Deep learning for edge computing applications: A state-of-the-art survey," *IEEE Access*, vol. 8, pp. 58322–58336, 2020.

[135] K. L.-M. Ang, J. K. P. Seng, E. Ngharamike, and G. K. Ijemaru, "Emerging technologies for smart cities' transportation: Geo-information, data analytics and machine learning approaches," *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 2, p. 85, Jan. 2022.

[136] M. H. Alkinani, A. A. Almazroi, M. Adhikari, and V. G. Menon, "Artificial intelligence-empowered logistic traffic management system using empirical intelligent XGBoost technique in vehicular edge networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4499–4508, Apr. 2023.

[137] H. Xu, Z. Cai, R. Li, and W. Li, "Efficient CityCam-to-edge cooperative learning for vehicle counting in ITS," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 16600–16611, Sep. 2022.

[138] Y. Shuai, J. Yimu, L. Shangdong, Y. Qifan, L. Kui, and S. Sisi, "MEMTV: A research on multi-level edge computing model for traffic video processing," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2020, pp. 3206–3210.

[139] J. Wang, J. Pan, and F. Esposito, "Elastic urban video surveillance system using edge computing," in *Proc. Workshop Smart Internet Things*, New York, NY, USA, Oct. 2017, pp. 1–6.

[140] Y.-Y. Li, T.-S. Lin, and H.-Y. Wei, "Game-theoretic cloud-edge resource allocation for video analytics in the factory of the future," in *Proc. IEEE VTS 17th Asia Pacific Wireless Commun. Symp. (APWCS)*, Aug. 2021, pp. 1–5.

[141] D. R. Patrikar and M. R. Parate, "Anomaly detection using edge computing in video surveillance system: Review," *Int. J. Multimedia Inf. Retr.*, vol. 11, no. 2, pp. 85–110, Mar. 2022.

[142] A. Manocha, G. Kumar, M. Bhatia, and A. Sharma, "Video-assisted smart health monitoring for affliction determination based on fog analytics," *J. Biomed. Informat.*, vol. 109, Sep. 2020, Art. no. 103513.

[143] T. Hussain, K. Muhammad, S. Khan, A. Ullah, M. Y. Lee, and S. W. Baik, "Intelligent baby behavior monitoring using embedded vision in IoT for smart healthcare centers," *J. Artif. Intell. Syst.*, vol. 1, no. 1, pp. 110–124, 2019.

[144] H. Abbas, F. Garberson, E. Glover, and D. P. Wall, "Machine learning for early detection of autism (and other conditions) using a parental questionnaire and home video screening," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 3558–3561.

[145] C. Taramasco, T. Rodenas, F. Martinez, P. Fuentes, R. Munoz, R. Olivares, V. H. C. De Albuquerque, and J. Demongeot, "A novel monitoring system for fall detection in older people," *IEEE Access*, vol. 6, pp. 43563–43574, 2018.

[146] W. Min, H. Cui, H. Rao, Z. Li, and L. Yao, "Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics," *IEEE Access*, vol. 6, pp. 9324–9335, 2018.

[147] Y.-Z. Hsieh and Y.-L. Jeng, "Development of home intelligent fall detection IoT system based on feedback optical flow convolutional neural network," *IEEE Access*, vol. 6, pp. 6048–6057, 2018.

[148] R. Rajavel, S. K. Ravichandran, K. Harimoorthy, P. Nagappan, and K. R. Gobichettipalayam, "IoT-based smart healthcare video surveillance system using edge computing," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 6, pp. 3195–3207, Jun. 2022.

[149] Q. Abbas, M. E. A. Ibrahim, and M. A. Jaffar, "Video scene analysis: An overview and challenges on deep learning algorithms," *Multimedia Tools Appl.*, vol. 77, no. 16, pp. 20415–20453, Aug. 2018.

[150] L. Hu and Q. Ni, "IoT-driven automated object detection algorithm for urban surveillance systems in smart cities," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 747–754, Apr. 2018.

[151] V. Tsakanikas and T. Dagiuklas, "Video surveillance systems-current status and future trends," *Comput. Electr. Eng.*, vol. 70, pp. 736–753, Aug. 2018.

[152] C. Long, Y. Cao, T. Jiang, and Q. Zhang, "Edge computing framework for cooperative video processing in multimedia IoT systems," *IEEE Trans. Multimedia*, vol. 20, no. 5, pp. 1126–1139, May 2018.

[153] S. Pirbhulal, O. W. Samuel, W. Wu, A. K. Sangaiah, and G. Li, "A joint resource-aware and medical data security framework for wearable healthcare systems," *Future Gener. Comput. Syst.*, vol. 95, pp. 382–391, Jun. 2019.

[154] Z. Xie, Y. Wang, and H. Wang, "Research on video surveillance system used for smart community," in *Proc. IEEE 5th Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2021, pp. 1947–1951.

[155] B. Kollar, L. Schneider, V. K. Horner, J. Zeller, M. Fricke, Z. Brugger, M. Gentz, J. Kiefer, and S. U. Eisenhardt, "Artificial intelligence-driven video analysis for novel outcome measures after smile reanimation surgery," *Facial Plastic Surg. Aesthetic Med.*, vol. 24, no. 2, pp. 117–123, Apr. 2022.

[156] S. Laha, R. Lalonde, A. E. Carmack, H. Foroosh, J. C. Olson, S. Shaikh, and U. Bagci, "Analysis of video retinal angiography with deep learning and Eulerian magnification," *Frontiers Comput. Sci.*, vol. 2, pp. 1–16, Jul. 2020.

[157] E. Kovalenko, A. Talitckii, A. Anikina, A. Shcherbak, O. Zimniakova, M. Semenov, E. Bril, D. V. Dylov, and A. Somov, "Distinguishing between Parkinson's disease and essential tremor through video analytics using machine learning: A pilot study," *IEEE Sensors J.*, vol. 21, no. 10, pp. 11916–11925, May 2021.

[158] S. Zheng, G. Liang, J. Chen, Q. Duan, and Y. Chang, "Severity assessment of cervical spondylotic myelopathy based on intelligent video analysis," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 9, pp. 4486–4496, Sep. 2022.

[159] M. R. Bhuiyan, J. Abdullah, N. Hashim, and F. Al Farid, "Video analytics using deep learning for crowd analysis: A review," *Multimedia Tools Appl.*, vol. 81, no. 19, pp. 27895–27922, Aug. 2022.

[160] M. Bendali-Braham, J. Weber, G. Forestier, L. Idoumghar, and P.-A. Müller, "Recent trends in crowd analysis: A review," *Mach. Learn. Appl.*, vol. 4, Jun. 2021, Art. no. 100023.

[161] X. Yang and H. Wang, "Design and implementation of intelligent analysis technology in sports video target and trajectory tracking algorithm," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, Jun. 2022, pp. 1–17.

[162] L. Zhao and D. Pan, "Video analysis of belt and road sports events based on wireless network and artificial intelligence technology," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, Feb. 2022, pp. 1–13.

[163] O. M. Teket and I. S. Yetik, "A fast deep learning based approach for basketball video analysis," in *Proc. 4th Int. Conf. Vis., Image Signal Process.*, New York, NY, USA, Mar. 2021, pp. 1–6.

[164] W. Guochen and S. Feihong, "Video analysis method of basketball training assistant based on deep learning theory during COVID-19 spread," *J. Intell. Fuzzy Syst.*, vol. 39, no. 6, pp. 8747–8755, Dec. 2020.

[165] R. Zhang, "Research on intelligent analysis technology of football video based on computer vision," in *Proc. 2nd Int. Conf. Internet Things Smart City*, May 2022, pp. 142–147.

[166] N. Liu, L. Liu, and Z. Sun, "Football game video analysis method with deep learning," *Comput. Intell. Neurosci.*, vol. 2022, Jun. 2022, Art. no. 3284156.

[167] G. Mackenzie and J. R. R. Ponte, "Video capture and analysis of cyclists using infrastructure in the act through machine learning," Centre Automot. Saf. Res., Univ. Adelaide, Adelaide, SA, Australia, Tech. Rep. CASR202, Sep. 2022.

[168] M. Zhou, G. H. Chen, P. Ferreira, and M. D. Smith, "Consumer behavior in the online classroom: Using video analytics and machine learning to understand the consumption of video courseware," *J. Marketing Res.*, vol. 58, no. 6, pp. 1079–1100, Dec. 2021.

[169] W. Mao, "Video analysis of intelligent teaching based on machine learning and virtual reality technology," *Neural Comput. Appl.*, vol. 34, no. 9, pp. 6603–6614, May 2022.

[170] B. Li, H. Zhang, and Y. Zhao, "Research based on video analysis the integration of the application of mathematics smart class teaching media and teachers' knowledge structure," in *Proc. Int. Joint Conf. Inf., Media Eng. (IJCIME)*, Dec. 2019, pp. 348–353.

[171] J. Stairs, R. Mangla, M. Chaudhery, J. S. Chandhok, and H. S. Timorabadi, "Engage AI: Leveraging video analytics for instructor-class awareness in virtual classroom settings," in *Proc. ASEE Virtual Annu. Conf. Content Access*, Jul. 2021, pp. 1–13.

[172] M. Rashmi, T. S. Ashwin, and R. M. R. Guddeti, "Surveillance video analysis for student action recognition and localization inside computer laboratories of a smart campus," *Multimedia Tools Appl.*, vol. 80, no. 2, pp. 2907–2929, Jan. 2021.

[173] K. Lv, Z. Sun, and M. Xu, "Artificial intelligent based video analysis on the teaching interaction patterns in classroom environment," *Int. J. Inf. Educ. Technol.*, vol. 11, no. 3, pp. 126–130, 2021.

[174] A. A. Mubarak, H. Cao, W. Zhang, and W. Zhang, "Visual analytics of video-clickstream data and prediction of learners' performance using deep learning models in MOOCs' courses," *Comput. Appl. Eng. Educ.*, vol. 29, no. 4, pp. 710–732, Jul. 2021.

[175] M. Shafique, A. Marchisio, R. V. W. Putra, and M. A. Hanif, "Towards energy-efficient and secure edge AI: A cross-layer framework," 2021, *arXiv:2109.09829*.

[176] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.

[177] K. Bonawitz, P. Kairouz, B. McMahan, and D. Ramage, "Federated learning and privacy," *Commun. ACM*, vol. 65, pp. 90–97, Mar. 2022.

[178] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.

ELARBI BADIDI received the bachelor's degree in electrical engineering and the M.Sc. degree in computer science from École Mohammedia des Ingénieurs, Rabat, Morocco, and the Ph.D. degree in computer science from Université de Montréal, Canada. He is currently an Associate Professor with the College of Information Technology (CIT), United Arab Emirates University (UAEU). Before joining UAEU, he worked for three years as a Bioinformatics Group Leader with the Biochemistry Department, Université de Montréal. He has over 15 years of research experience in service-oriented architecture, cloud computing, and context-aware systems, focusing on quality of service management, service level agreement management, quality of context (QoC) negotiation, QoC-based

selection, and data-as-a-service provisioning. He co-edited a handbook on smart cities and published over 80 peer-reviewed papers in reputed international journals and conferences and ten book chapters. He served on the technical program committees of many international conferences and as a reviewer of several journals. His current research interests include edge computing, cloud computing, the Internet of Things (IoT), big data, data stream processing, and data analytics.

KARIMA MOUMANE received the engineering degree in computer science from ENSIAS, Rabat, in 2011, and the Ph.D. (Eng.) degree in computer engineering from ENSIAS, Mohammed V University, in March 2018. She has eight years of experience as a Software Engineer with the Moroccan Customs (ADII), from 2011 to 2019, and one and a half years as an Accredited Trainer with the Moroccan Customs Training Institute (IFD), from 2019 to 2020. She has been a Professor/Researcher with ENSIAS, Mohammed V University, Rabat, since December 2020. She obtained a number of certificates in different computer science disciplines, such as the Scrum Master Certificate, ITIL®V4 Foundation Certificate, and ISTQB Certified Tester Foundation Level. Recently, she has worked on several papers submitted to indexed journals and conferences, most of which are related to the application of machine learning techniques to different domains (healthcare area, supply chain, oceanography, environment management, and renewable energy), software quality evaluation of m-health apps, and video analytics for smart cities.

FIRDAOUS EL GHAZI received the Graduate degree from the Mohammadia Engineering School (EMI), Morocco, and the Ph.D. degree in energetics from Ibn Tofail University, Morocco, in 2021. She is currently an industrial engineer. She specializes in supply chain and has 12 years of experience in the energy sector. She is also working on energy-related themes in order to support the dynamics of energy transition in Morocco and African countries.

• • •