

Edge AI for Real-Time Video Analytics in Surveillance Systems

Dr K P N V Satyasree¹, Dr. Taviti Naidu Gongada², Niraj Upadhayaya³, Naresh E⁴, Dr. Jyoti Prasad Patra⁵, Dr. Mandeep Kaur⁶

¹Usha Rama College of Engineering and Technology, Telaprolu
satyasreekpnv@gmail.com

²GITAM School of Business,
GITAM (Deemed to be) UNIVERSITY, VISAKHAPATNAM
tgongada2@gitam.edu

³SRM University AP, Amravati, Andhra Pradesh 522510
nirajup@gmail.com

⁴Manipal Institute of Technology, Bengaluru
naresh.e@manipal.edu

⁵Krupaja Engineering College, Bhubaneswar 751002
jpp42003@yahoo.co.in

⁶Department of Electronics and Communication Engineering, Punjabi University, Punjab 147002
ermandeep0@gmail.com

Abstract: More and more surveillance systems are being used to increase security and safety for the general public. However, the conventional method of processing all video data in the cloud can be ineffective and slow response times. By performing video analytics at the network's edge, close to where the data is created, Edge AI is a promising new strategy that can address these issues.

The most recent developments in edge AI for real-time video analytics in security systems are discussed in this paper. We discuss the different techniques that are being used, as well as the applications that are being enabled by edge AI. The paper also discusses the challenges and limitations of edge AI, and the future research directions in this area.

Keywords: Edge AI, Surveillance Systems, Real-Time Video Analytics, Object Detection, Anomaly Detection, Facial Recognition, Edge Computing.

I. Introduction:

Modern security infrastructure is incomplete without surveillance systems, which act as watchful sentinels to protect our cities, transportation systems, commercial buildings, and critical infrastructure. These systems produce a massive flood of video data that needs to be quickly and accurately analyzed in order to identify security threats, track crucial operations, and effectively handle emergencies. Video analytics in surveillance has historically relied heavily on cloud-based solutions, which, while effective, pose significant difficulties.[7]; These difficulties include issues with data privacy and security, high bandwidth consumption, and latency in data transmission. In response to these limitations, a revolutionary shift in the surveillance industry is in progress, as evidenced by the growing use of Edge AI, in which artificial intelligence algorithms are built right into edge devices like cameras and network video recorders.[1]; With this change, real-time video analytics enters a new era that promises to transform how we think about surveillance by bringing computation closer to the data source. Fig 1 explains machine learning in video surveillance. Edge AI offers a novel method for the analysis of video data because it can carry out complex computations locally on edge devices.

Surveillance systems can benefit greatly from utilizing the computational power of edge devices and integrating AI models right inside of them.[2]; One of these is a significant decrease in latency, which guarantees that crucial insights are generated in real-time and enables quicker response times to security incidents.[8]; Furthermore, Edge AI solutions significantly lessen the need for cloud resources, reducing bandwidth snarls and providing more affordable and scalable security systems.[3]; Deploying AI algorithms at the edge also improves data privacy because private video data can be processed there rather than being sent to distant servers, all while addressing urgent privacy concerns.

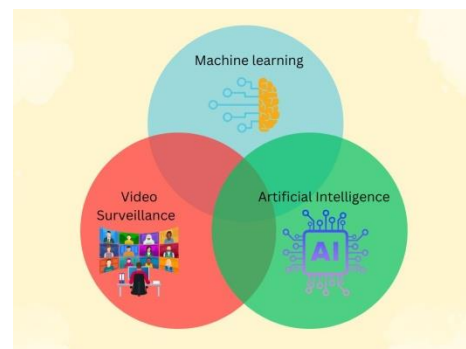


Fig 1 :Machine Learning in video Surveillance

In the context of surveillance systems, this research paper aims to examine the complex interaction between edge AI and real-time video analytics. It explores the technical foundations, benefits, and difficulties of this paradigm shift with an emphasis on its potential for transformation.[4]; We will analyze the approaches, designs, and case studies that best demonstrate the use of edge AI in surveillance throughout this paper. We will also explore current trends and promising new directions, providing insights into how Edge AI can develop to meet the constantly evolving needs of modern surveillance applications.[5]; We will explore the fascinating nexus between technology and security as we explore the promising horizons of edge AI, where the fusion of intelligence and immediacy creates a new era of surveillance systems, fostering safer and more secure communities.

II. Literature Review:

Real-time video analytics are now more important than ever because surveillance systems are becoming more sophisticated and widely used. The incorporation of edge AI technology has emerged as a game changer for real-time video analytics in surveillance systems. It is now simpler and more effective to process video data in a decentralized network for use in a surveillance system due to the increased use of AI technologies and IoT devices (Yu et al., 2020).

Edge AI is used in surveillance systems to process video data intelligently at the front-end camera, improving the system's capacity to recognize and respond to security threats in real-time.

This strategy makes use of edge computing, which allows data to be processed and analyzed locally at edge devices rather than being sent to a central server. This has been shown to be more effective and efficient because it lessens the load on computers, networks, and data transmission. Edge computing also makes it possible to make decisions and respond quickly by reducing latency and ensuring real-time analysis. The video surveillance system service platform's capacity for intelligent processing can be greatly enhanced with the use of edge AI.

Sergio Saponara, Abdussalam Elhanashi, Alessio Gagliardi(2020) Real-time video fire/smoke detection based on CNN in antifire surveillance systems

This author explains a real-time, embedded implementation of a fire and smoke detection method that makes use of common security cameras. The goal of this work is to create intelligent Internet of Things (IoT) devices for indoor and outdoor fire/smoke detection. For real-time fire/smoke detection, the proposed method utilizes YOLOv2 Convolution Neural Network, which outperforms other cutting-edge techniques. The model was successfully implemented in a low-cost embedded device (Jetson Nano) after being trained on a large

scale of fire/smoke and negative videos in various indoor and outdoor scenarios.

LCDnet: a lightweight crowd density estimation model for real-time

video surveillance(2022) et.al the author introduces the LCDnet model, a new, simple crowd density estimation model for in-the-moment video surveillance. The authors evaluate LCDnet's performance using several benchmark datasets and suggest a better training method based on curriculum learning. The findings demonstrate that LCDnet outperforms existing models in accuracy while significantly reducing inference time and memory requirements. For researchers and practitioners interested in crowd counting and density estimation, the paper provides a thorough description of the proposed model and its evaluation.

Edge AI in Surveillance Systems:

1. Introduction to Edge AI:

Edge AI is the practice of installing artificial intelligence (AI) algorithms directly on edge hardware, such as cameras, sensors, and network video recorders (NVRs), as opposed to relying solely on distant cloud-based servers to process data. Real-time data analysis is made possible by this method at or close to the data collection source. Edge AI, also referred to as edge artificial intelligence, is the deployment of artificial intelligence algorithms and models directly on edge devices like smart phones, IoT devices, and embedded systems. This is in contrast to using centralized cloud servers for processing. This decentralized approach allows for real-time, low-latency data analysis and decision-making at the network's edge, which eliminates the need for constant internet connectivity and addresses privacy and security concerns. Edge AI enables devices to carry out operations like image recognition, natural language processing, and predictive analytics locally in sectors like autonomous vehicles, industrial automation, smart cities, and healthcare, where rapid response times and data privacy are essential.

2. Advantages of Edge AI in Surveillance:

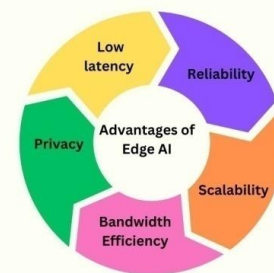


Fig 2: Advantages of Edge AI

Low Latency: One of the primary advantages of Edge AI in surveillance is reduced latency. By processing video data locally on edge devices, the time between data capture and analysis is minimized. This low latency is critical for real-time threat detection and rapid response in security applications. Low latency in the context of Edge AI refers to the shortest possible delay in data processing and generating AI-driven insights directly on edge devices, such as cameras or IoT devices, without relying on distant cloud servers for analysis. The processing of data in real-time or nearly real-time at the edge enables quick response to events and anomalies. In applications like autonomous vehicles, industrial automation, and surveillance, where split-second decisions can have an impact on safety, security, and operational efficiency, low latency in Edge AI is especially important. Edge AI ensures that crucial actions can be taken quickly and increases the overall effectiveness of AI-powered systems while reducing reliance on external networks and cloud resources by speeding up local data analysis.

Bandwidth Efficiency: The ability of AI to efficiently use bandwidth by minimizing the amount of data that needs to be transmitted across networks optimizes data usage. AI algorithms, especially those used at the network's edge, can locally preprocess data, removing only relevant insights or anomalies. This reduces the strain on network infrastructure, conserves bandwidth, and brings down the price of data transfer. AI-driven bandwidth efficiency improves user experience by reducing delays and congestion in applications like remote monitoring, IoT devices, and video streaming, where data transmission is essential. It also contributes to more sustainable and economical data management, making it an invaluable asset in our increasingly connected world.

Privacy and Security: Edge AI involves processing sensitive data directly on edge devices, so privacy and security are top priorities. This decentralized strategy improves data privacy and security by minimizing data exposure during transmission and lowering vulnerability to cyberattacks. Edge AI systems make it easier to comply with data privacy laws and can incorporate strong encryption, authentication, and real-time threat detection mechanisms. To reduce vulnerabilities and guarantee long-term security, edge devices must maintain regular updates and patch management. Edge AI ultimately strikes a balance between data-driven insights and protecting personal information while bolstering security measures for a variety of applications, from autonomous vehicles to IoT devices and beyond.

Scalability: The ability to effectively grow and modify edge computing environments in order to meet changing AI requirements and growing data volumes is referred to as scalability in edge AI. Edge AI systems should be flexible,

allowing for the simple integration of additional edge devices and the flexibility to change processing power as necessary. For applications like smart cities, industrial automation, and IoT, where the number of edge devices can rise quickly, this scalability is essential. Organizations can harness the power of distributed computing and artificial intelligence while maintaining responsiveness and efficiency in dynamic and evolving edge environments by ensuring that Edge AI solutions can scale seamlessly.

Reliability: Any system or technology must be reliable in order to ensure consistent performance and uptime under a variety of circumstances. Since Edge AI involves implementing artificial intelligence algorithms on edge devices like IoT sensors or autonomous systems, reliability is essential. It is crucial that these devices function continuously because they frequently operate in isolated or challenging environments. A dependable Edge AI system can continue to carry out its tasks without jeopardizing security or effectiveness despite hardware malfunctions, network outages, and power outages. Reliability ensures that crucial decisions and operations relying on AI continue without interruption, whether in autonomous vehicles, industrial automation, or healthcare, ultimately enhancing the overall credibility and efficacy of edge-based AI applications.



Fig 3: DNN in video analytics

Proposed method:

In Edge AI for Real-Time Video Analytics in Surveillance Systems, Deep Learning models play a key role in enabling the extraction of valuable insights from video streams at the edge. In order to accurately identify objects and people of interest in real-time, Convolution Neural Networks (CNNs) are essential for object detection, facial recognition, and scene analysis. Fig 3 explains DNN in video analytics. Recurrent neural networks (RNNs) are crucial for behavior recognition and anomaly detection because they are excellent at tracking and analyzing temporal patterns. Comprehensive video analysis is made possible by hybrid models that incorporate CNN and RNN components, while 3D Convolutional Networks and Two-Stream Networks improve the spatiotemporal comprehension of video data. These deep learning models give surveillance systems the autonomy they

need to operate at the edge while providing quick, context-sensitive insights that are essential for boosting security and operational effectiveness.

2.2 Deep Learning Models:

Deep learning models are a fundamental component of Edge AI in surveillance. These models are responsible for tasks like object detection, facial recognition, and behavior analysis. Some common models include:

Convolution Neural Networks (CNNs):

Convolution neural networks (CNNs) are heavily utilized by Edge AI for real-time video analytics in security systems. These deep learning models excel at extracting complex features from video frames to accurately detect, track, and recognize objects, like faces and license plates. Due to their capacity to process data in parallel, which is frequently accelerated by specialized hardware, they are crucial for quick event detection and response. Using CNNs at the edge lessens reliance on the cloud while enhancing privacy and bandwidth efficiency. Fig 4 explains convolution layer in video frame surveillance Continuous improvements in real-time video analytics, made possible by improvements in CNN architectures, hardware acceleration, and edge computing capabilities, increase the efficacy of surveillance systems. One category of deep learning models that is particularly effective for image and video analysis is convolution neural networks. They include convolution layers, pooling layers, and fully connected layers, among others. In surveillance systems, CNNs are commonly used for tasks such as object detection, facial recognition, and scene classification. Here's a breakdown of key elements:

- **Convolution Layers:** These layers apply convolution operations to the input image, using small filters or kernels to detect features like edges, corners, and textures. Convolutional layers learn to recognize low-level visual patterns.
- **Pooling Layers:** Pooling layers down sample the spatial dimensions of feature maps, reducing computational complexity while retaining important information. Max-pooling and average-pooling are common pooling techniques.
- **Fully Connected Layers:** In the final layers of a CNN, fully connected layers process the high-level features extracted by earlier layers and make predictions based on them. For instance, in facial recognition, these layers might determine whether a detected face matches a known individual.

- **Architectures:** Various CNN architectures have been developed for specific tasks. For instance, YOLO (You Only Look Once) and SSD (Single Shot Multi Box Detector) are designed for real-time object detection. VGG Net and Res Net are popular choices for image classification, and Siamese networks are used for one-shot learning tasks.

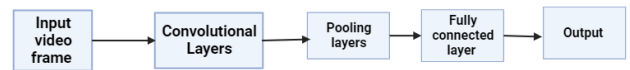


Fig 4: Convolution layer in video Surveillance

Recurrent neural networks are useful in surveillance systems for analyzing patterns over time because they are built to handle sequential data. RNNs can be used in video analytics for tasks like object tracking and spotting unusual behaviors. RNNs' essential elements include:

Recurrent Layers: These layers keep track of previous time steps in a sequence in a hidden state. RNNs can now simulate the temporal dependencies in data.

LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit): The vanishing gradient issue is solved by the specialized RNN architectures LSTM and GRU, which enable RNNs to detect distant dependencies in sequences.

RNN variants that process sequences both forward and backward are known as bidirectional RNNs, and they enhance the modeling of context in surveillance tasks.

Applications include identifying particular actions or behaviors in video streams, tracking the movement of objects or people across frames, and anomaly detection based on temporal patterns.

Recurrent Neural Networks (RNNs): By offering temporal context and sequence analysis capabilities, recurrent neural networks (RNNs) play a crucial role in Edge AI for real-time video analytics in surveillance systems. RNNs can model dynamic patterns over time, making them useful for tasks like video-based anomaly detection, behavior recognition, and tracking in surveillance, whereas CNNs excel at static frame analysis. RNNs improve the contextual understanding of events by processing sequential data

efficiently on edge devices, enabling the detection of complex scenarios and shady activities. In real-time video analytics at the edge, the combination of CNNs for frame-level analysis and RNNs for temporal modeling enables surveillance systems to not only recognize objects and people but also to interpret their interactions and behavior. Fig 5 explains deep learning in edge system. These layers keep track of previous time steps in a sequence in a hidden state. RNNs can now simulate the temporal dependencies in data. LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit): The vanishing gradient issue is solved by the specialized RNN architectures LSTM and GRU, which enable RNNs to detect distant dependencies in sequences. RNN variants that process sequences both forward and backward are known as bidirectional RNNs, and they enhance the modeling of context in surveillance tasks.

Applications include locating specific actions or behaviors in video streams, following objects or people as they move between frames, and detecting anomalies based on temporal patterns. Edge AI is a subset of artificial intelligence (AI) that works close to where data is produced, at the network's edge. In contrast, cloud-based computations are used in traditional AI. Real-time video analytics are being performed in surveillance systems using edge AI. This can be applied to enhance traffic management, identify suspicious activity, and take appropriate action.

Edge AI is a cutting-edge innovation that has the potential to transform security systems.

Anomaly detection, facial recognition, object tracking, and other edge AI techniques are employed in surveillance systems. Deep learning models for image analysis like Convolution Neural Networks (CNNs) are frequently used to implement these techniques.

Real-time video analytics are carried out by surveillance systems using edge AI, a type of artificial intelligence. With the aid of this technology, traffic can be managed and necessary action taken after spotting suspicious activity. The lack of computational power on edge devices, the requirement to optimize models for efficient inference, and the maintenance of system dependability under various environmental factors are all difficulties. It is important to carefully weigh the trade-offs between local processing and cloud-based analysis. Edge AI architectures include

lightweight models for edge deployment, optimized software frameworks, and hardware accelerators like GPUs and TPUs to accelerate AI inference. In spite of these difficulties, edge AI has the power to transform security systems by improving traffic management, spotting suspicious activity, and taking the appropriate action.

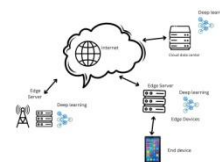


Fig 5: Deep learning in edge system

Deep neural networks (DNNs) are a subclass of machine learning algorithms that can be applied to a range of tasks, such as object detection, natural language processing, and image classification. DNNs can learn to recognize patterns in the data that are hidden to humans by being trained on large datasets of labeled data.

Artificial intelligence (AI) known as "edge AI" is used on edge devices like cameras, sensors, and gateways. Real-time data processing is made possible by edge AI, which is crucial for many applications, including video analytics in security systems.

III. Performance metrics:

Convolution layer performance metrics are essential for evaluating how well these layers perform at analyzing and processing video data in a video surveillance system. Accuracy, which gauges the proportion of correctly classified objects or events, speed, which assesses the computational effectiveness of the convolution operations, and detection rate, which measures the system's aptitude to recognize pertinent objects or events within the video stream, are some of the frequently used metrics. table 1 shows performance metrics value of CNN. Furthermore, important metrics for assessing the trade-off between false positives and false negatives include precision and recall. The F1-score can offer a balance between recall and precision. Furthermore, latency, which measures the time required to process each frame and ensures prompt responses to potential security incidents, becomes a crucial metric in real-time surveillance applications.

Table 1: Results of video surveillance using CNNs

Videos	Accuracy (%)	True positive	True Negative
Video 1	92.3	245	198
Video 2	89.8	987	145
Video 3	93.6	564	296

IV. Result and discussion:

In a video surveillance system, performance metrics such as True Positive (TP) and True Negative (TN) are often used to assess the accuracy of object detection and recognition tasks performed by convolutional neural networks (CNNs) or similar models. These metrics are typically used in the context of binary classification problems, where the goal is to distinguish between two classes, such as "object present" and "object absent" in the context of surveillance. Here's how you can interpret these metrics and their values in a performance metrics table:

1. True Positive (TP): This metric represents the number of objects correctly detected as being present in the video frames. In the context of surveillance, this would mean instances where the system correctly identifies and reports the presence of an object of interest (e.g., a person or a vehicle).
2. True Negative (TN): This metric represents the number of instances where the system correctly identifies the absence of the object of interest. In surveillance, this would be situations where the system correctly recognizes that there is no object of interest in the frame.

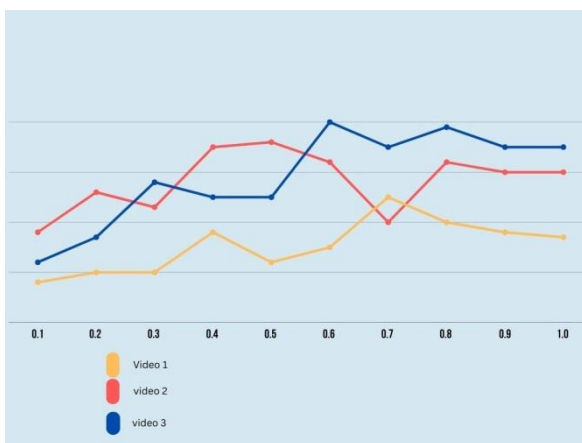


Fig 5: Videos true Positive value in percentage

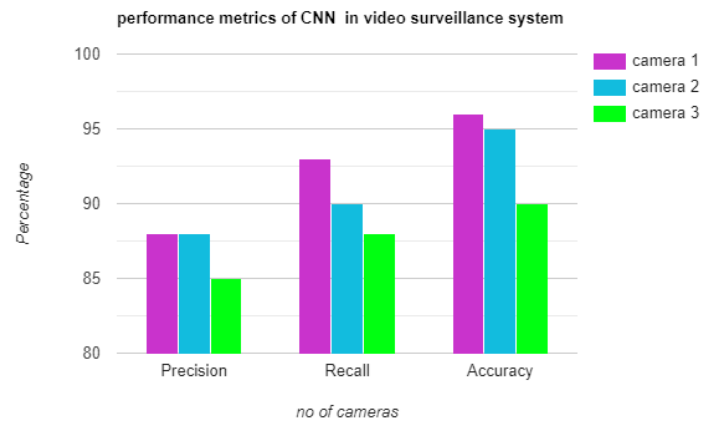


Fig 6: Performance metrics of CNN in video surveillance system

V. Challenges and Considerations:

Both opportunities and challenges are presented by the use of Edge AI for real-time video analytics in surveillance systems. The processing of sensitive visual data on-device raises privacy issues that must be addressed as well as the need for powerful edge computing hardware to handle high-resolution video streams in real-time. It can be difficult to scale and maintain distributed edge devices across a surveillance network. The benefits of local video analysis, on the other hand, lie in the opportunities for improved security and decreased latency, which reduce the need for data transmission to the cloud. Edge AI offers a complete security solution by enabling customization to particular surveillance tasks and allowing for seamless integration with other sensors. Furthermore, real-time video analytics in surveillance systems have the potential to be fully realized thanks to ongoing hardware and machine learning advancements that promise to solve current problems.

VI. Future Trends and Research Directions:

Research in Edge AI for surveillance continues to evolve. Future trends may include federated learning for collaborative analysis, integration with 5G networks for enhanced data transfer, and improved power-efficient hardware for edge devices. Additionally, research efforts will focus on addressing privacy concerns and ensuring robustness against adversarial attacks. Edge AI in surveillance systems offers a powerful solution to the growing demands of real-time video analytics, enhancing security, reducing latency, and ensuring data privacy. As technology advances, it is poised to play an increasingly central role in modern security infrastructure.

References:

1. Czyewski, Andrzej & Szwoch, Grzegorz & Dalka, Piotr & Szczuko, Piotr & Ciarkowski, Andrzej & Ellwart, Damian & Merta, Tomasz & Łopatka, Kuba & Kulasek, ukasz & Wolski, J?drzej. (2011). Multi-Stage Video Analysis Framework. 10.5772/16088.
2. Yu, K., Qi, X., Sato, T., Myint, S H., Wen, Z., Katsuyama, Y., Tokuda, K., Kameyama, W., & Sato, T. (2020, January 1). Design and Performance Evaluation of an AI-Based W-Band Suspicious Object Detection System for Moving Persons in the IoT Paradigm. <https://scite.ai/reports/10.1109/access.2020.2991225>
3. Call for Papers: Special Issue on Real-time Video Surveillance/Security Systems. (2000, April). *Real-Time Imaging*, 6(2), 173. [https://doi.org/10.1016/s1077-2014\(00\)90206-7](https://doi.org/10.1016/s1077-2014(00)90206-7)
4. Lyu, Z., & Luo, J. (2022, October 9). A Surveillance Video Real-Time Object Detection System Based on Edge-Cloud Cooperation in Airport Apron. *Applied Sciences*, 12(19), 10128. <https://doi.org/10.3390/app121910128>
5. Gaikwad, B., & Karmakar, A. (2022, April). End-to-end person re-identification: Real-time video surveillance over edge-cloud environment. *Computers and Electrical Engineering*, 99, 107824. <https://doi.org/10.1016/j.compeleceng.2022.107824>
6. Cooharajanane, N., Kasamwattanarote, S., Lipikorn, R., & Satoh, S. (2012, October 20). Automated real-time video surveillance summarization framework. *Journal of Real-Time Image Processing*, 10(3), 513–532. <https://doi.org/10.1007/s11554-012-0280-7>
7. Diamantopoulos, G., & Spann, M. (2005, June). Event detection for intelligent car park video surveillance. *Real-Time Imaging*, 11(3), 233–243. <https://doi.org/10.1016/j.rti.2005.02.002>
8. Foresti, G. (1998). A real-time system for video surveillance of unattended outdoor environments. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(6), 697–704. <https://doi.org/10.1109/76.728411>
9. Pázsit, I. (2004, May). Diagnostics and Surveillance Methods in Nuclear Systems for Real-Time Applications. *Real-Time Systems*, 27(1), 97–113. <https://doi.org/10.1023/b:time.0000019129.88316.c7>
10. Manogaran, G., Baskar, S., Shakeel, P. M., Chilamkurti, N., & Kumar, R. (2019, May 10). Analytics in real time surveillance video using two-bit transform accelerative regressive frame check. *Multimedia Tools and Applications*, 79(23–24), 16155–16172. <https://doi.org/10.1007/s11042-019-7526-3>
11. Singh, S., Shekhar, C., & Vohra, A. (2016, March 11). FPGA-Based Real-Time Motion Detection for Automated Video Surveillance Systems. *Electronics*, 5(4), 10. <https://doi.org/10.3390/electronics5010010>
12. Saponara, S., Elhanashi, A., & Gagliardi, A. (2020, November 10). Real-time video fire/smoke detection based on CNN in antifire surveillance systems. *Journal of Real-Time Image Processing*, 18(3), 889–900. <https://doi.org/10.1007/s11554-020-01044-0>
13. Khan, M. A., Menouar, H., & Hamila, R. (2023, March 6). LCDnet: a lightweight crowd density estimation model for real-time video surveillance. *Journal of Real-Time Image Processing*, 20(2). <https://doi.org/10.1007/s11554-023-01286-8>
14. Holstein, K., McLaren, B. M., & Alevin, V. (2019, July 22). Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity. *Journal of Learning Analytics*, 6(2). <https://doi.org/10.18608/jla.2019.62.3>
15. Video analytics under surveillance. (2009, January 23). *Sensor Review*, 29(1). <https://doi.org/10.1108/sr.2009.08729aab.002>
16. Gaikwad, B., & Karmakar, A. (2021, February 2). Smart surveillance system for real-time multi-person multi-camera tracking at the edge. *Journal of Real-Time Image Processing*, 18(6), 1993–2007. <https://doi.org/10.1007/s11554-020-01066-8>
17. Chen, Jiasi & Ran, Xukan. (2019). Deep Learning With Edge Computing: A Review. Proceedings of the IEEE. PP. 1–20. 10.1109/JPROC.2019.2921977.