

Smart CCTV Surveillance using Edge AI and Blockchain: A Review

Pawar Sakshi Ajinath¹, Bangar Abhishek Popat², Bhosale Sachin Bajirao³, Khatri Anand A.⁴

¹Department of Computer Engineering / Jaihind College of Engineering / SPPU, Pune, India

Abstract- The rapid expansion of smart cities and urban populations has created an urgent demand for real-time, highly secure, and privacy-preserving public surveillance systems. Traditional CCTV surveillance architectures heavily rely on centralized cloud servers, making them vulnerable to single points of failure, network latency, data tampering, and severe privacy breaches. To address these critical limitations, this paper proposes a secure, decentralized surveillance framework that integrates Edge AI with private blockchain technology and InterPlanetary File System (IPFS). The proposed system utilizes lightweight deep learning models—such as YOLOv8 for high-precision real-time object detection, DeepSORT for multi-object tracking, and ArcFace for robust facial recognition—operating locally on edge devices to achieve low-latency processing without relying on centralized cloud infrastructure. To ensure strict data confidentiality and integrity, a hybrid cryptography model combining AES-256 for video metadata encryption and RSA-2048 for secure key management within a Trusted Execution Environment (TEE) is implemented. Furthermore, a private Ethereum blockchain leveraging Quorum Byzantine Fault Tolerance (QBFT) consensus and smart contracts is deployed to enforce role-based access control (RBAC) and maintain tamper-proof, immutable audit trails. Experimental results demonstrate that the proposed edgeblockchain integration significantly improves detection precision, ensures real-time performance at 60 FPS, and provides robust resistance against data tampering, offering an efficient and secure solution for next-generation intelligent urban surveillance.

Keywords— Edge AI; CCTV Surveillance; Private Blockchain; IPFS; YOLOv8; DeepSORT; Hybrid Cryptography; Quorum Byzantine Fault Tolerance (QBFT); Smart Contracts

I. INTRODUCTION

The rapid growth of smart cities and urban populations has created an urgent demand for reliable public safety and surveillance solutions. Traditional CCTV surveillance systems rely on centralized cloud architectures, which face limitations in monitoring dynamic environments and responding to threats in real time. Furthermore, centralized cloud storage creates critical vulnerabilities, including single points of failure, network bandwidth bottlenecks, and risks of data tampering.

To address these challenges, integrating Edge Artificial Intelligence (Edge AI) with decentralized technology has become a transformative approach. Edge AI uses deep learning models—such as YOLOv8 for object detection, DeepSORT for tracking, and ArcFace for facial recognition—to run inference locally on edge devices like Raspberry Pi or NVIDIA

Jetson, reducing latency and cloud reliance. However, edge systems alone still struggle with data integrity and secure evidence management.

To overcome these gaps, this study proposes a decentralized CCTV surveillance framework that integrates Edge AI with a private Ethereum blockchain, InterPlanetary File System (IPFS), and hybrid cryptography (AES-256 and RSA-2048). In this system, edge devices perform real-time threat detection and offload encrypted video evidence to IPFS for distributed storage. Simultaneously, immutable metadata and cryptographic hashes (CIDs) are anchored onto a private blockchain using Quorum Byzantine Fault Tolerance (QBFT) consensus and smart contracts to enforce Role-Based Access Control (RBAC). The proposed framework aims to eliminate single points of failure, ensure tamper-proof legal evidence, and provide a secure solution for smart city security.

II. REVIEW METHODOLOGY

A systematic review of selected research papers was conducted to understand recent developments in intelligent surveillance systems, edge computing, decentralized blockchain architectures, and deep learning-based threat detection. The selected studies were analyzed based on their detection techniques, deep learning models, hardware deployment platforms, cryptographic security mechanisms, performance metrics, and identified limitations. The reviewed approaches include YOLO-based object detection, multi-object tracking using DeepSORT, facial recognition models like ArcFace, Error Level Analysis (ELA), hybrid cryptography, and private blockchain integration with InterPlanetary File System (IPFS).

The reviewed studies show that YOLO-based architectures are highly effective for realtime object detection and tracking in dynamic environments. Furthermore, edgecomputing frameworks deployed on low-power devices minimize latency and reduce cloud dependency, while decentralized blockchain technologies and IPFS ensure tamper-proof metadata storage and data integrity. Hybrid cryptographic models combining symmetric and asymmetric encryption secure sensitive video footage and management keys during transmission.

III. LITERATURE REVIEW

1. Edge-Based Object Detection and Tracking

Edge computing has significantly transformed modern smart surveillance architectures by shifting intensive AI inference and decision-making capabilities from traditional centralized cloud servers directly to local cameras and edge devices (such as Raspberry Pi or NVIDIA Jetson platforms). In conventional setups, continuous streaming of raw high-definition video to cloud servers consumes massive network bandwidth, creates significant processing latency, and exposes sensitive video streams to potential network interception. Edge-AI mitigates these core limitations by processing visual data locally at the source and transmitting only critical event metadata or encrypted threat logs. Recent literature highlights the effectiveness of

lightweight deep learning pipelines for ondevice threat analysis. For instance, Galhe et al. demonstrated that combining intelligent CCTV networks with edge processing enables real-time automated threat detection—such as recognizing concealed weapons or unauthorized entry—achieving high precision (up to 95%) and immediate alert dispatching. Building on localized inference, Nandankar and Khonde evaluated lightweight vision pipelines combining YOLOv8 for high-speed object detection and DeepSORT for multi-object spatialtemporal tracking on low-power edge nodes. By maintaining unique object IDs across consecutive video frames locally, these architectures eliminate frame dropping and reduce processing latency to milliseconds. Furthermore, integrating facial recognition algorithms like ArcFace alongside edge object detectors allows the system to differentiate authorized personnel from unknown intruders without exposing biometric databases to external cloud networks. This local execution drastically reduces network bandwidth requirements, ensures operational continuity even during internet downtime, and provides a privacy-preserving foundation for next-generation smart city security systems

2. Blockchain, IPFS, and Data Security

Centralized storage models are vulnerable to insider threats and data manipulation. To address this, recent research by Sarar et al. demonstrates a decentralized approach integrating private blockchain technology with the InterPlanetary File System (IPFS) for secure evidence management. To avoid the high storage costs of saving large video files directly on a blockchain, an off-chain storage model is used. Encrypted video payloads are uploaded to IPFS, which generates a unique cryptographic hash known as a Content Identifier (CID). Any attempt to alter the video alters this hash, immediately flagging tampering. These CIDs, along with metadata such as camera IDs and timestamps, are anchored onto a private Ethereum blockchain using Quorum Byzantine Fault Tolerance (QBFT) consensus. Selfexecuting smart contracts enforce Role-Based Access Control (RBAC), ensuring that only authorized personnel can access and inspect

the audit trails, maintaining a strict chain of custody for legal evidence.

3. Hybrid Cryptography and Key Management

To guarantee end-to-end data confidentiality and privacy protection, hybrid encryption mechanisms are applied directly at the edge layer before storing or transmitting surveillance footage across untrusted networks. Standalone cryptographic algorithms often fail to strike a balance between computational efficiency and high-level security; symmetric ciphers are fast but face key distribution challenges, while asymmetric ciphers offer secure key exchange at the cost of high processing overhead.

To solve this, hybrid cryptography combines the strengths of both approaches. Advanced Encryption Standard (AES-256) in Galois/Counter Mode (GCM) is utilized at the edge node to rapidly encrypt high-volume raw video payloads and event metadata with minimal computational latency. Simultaneously, RSA-2048 (or Elliptic Curve Cryptography, ECC) is employed to securely wrap and distribute the generated symmetric AES keys to authorized network participants. To safeguard cryptographic keys from sophisticated insider threats, memory extraction, or unauthorized system administrator overrides, key management and decryption routines are isolated off-chain inside a hardware-backed Trusted Execution Environment (TEE), such as Intel SGX or ARM TrustZone. The Key Management System (KMS) leverages these secure enclaves to process decryption requests only after validating smart contract access permissions. This multi-layered cryptographic security stack ensures end-to-end confidentiality, prevents unauthorized data exposure during transit and storage, and preserves the legal chain-of-custody required for courtadmissible video evidence.

4. AI-Powered Traffic and Violation Detection

Beyond basic physical threat detection, recent literature explores the application of Edge AI in automated urban traffic monitoring and dynamic violation detection. Intelligent transportation systems increasingly rely on high-speed camera networks to regulate traffic flow and enforce road safety compliance automatically. Research by Polu et

al. demonstrated that continuous CCTV video surveillance integrated with YOLO-based object detection, Optical Flow algorithms, and spatiotemporal anomaly recognition models can automatically identify dynamic infractions—such as overspeeding, wrong-way driving, illegal lane changes, and traffic signal violations—in real time with high operational reliability.

To process dense urban traffic streams efficiently without overloading central municipal networks, lightweight inference engines running directly on roadside edge units (such as smart traffic light controllers or localized roadside units) filter non-event video frames locally. When a violation is detected, the system automatically captures high-resolution vehicle license plate snapshots using Automatic Number Plate Recognition (ANPR) algorithms.

To ensure legal compliance and address cross-jurisdictional data privacy concerns across distributed city surveillance nodes, these architectures incorporate federated learning alongside blockchain-backed evidence logging. Federated learning allows edge cameras to collaboratively train and improve anomaly detection models without transmitting sensitive, raw vehicle video feeds to a centralized server. Concurrently, immutable records of traffic violations—including vehicle plate hashes, location telemetry, and time stamped video clips—are cryptographically anchored to a distributed ledger. This multi-node setup ensures that urban violation records remain untampered and legally admissible while enabling smart city traffic management infrastructure to respond to road anomalies dynamically. GVLM, the full model reached 99.39% overall accuracy, 95.36% F1-score, and 90.02% IoU, compared with 97.16%, 89.11%, and 84.56% for the baseline without the two components.

5. Summary and Research Need

Overall, the reviewed literature demonstrates that combining Edge AI (such as YOLOv8 and DeepSORT) with decentralized technologies—specifically private Ethereum blockchains, IPFS, and hybrid cryptography—effectively enables real-time threat

detection and secure evidence management. Furthermore, mechanisms like Role-Based Access Control (RBAC) ensure strong privacy compliance. However, existing studies reveal notable technical gaps. Executing intensive deep learning inference alongside multi-layered encryption on lowpower edge nodes introduces high computational overhead and processing bottlenecks. Additionally, current works often treat edge performance and blockchain security in isolation, lacking a unified end-to-end architecture, while offloading encrypted data to IPFS can introduce latency during peak conditions. These identified gaps highlight the critical need for an optimized, low-latency surveillance framework that harmonizes real-time

edge processing with robust decentralized security, forming the primary motivation of this study.

IV. COMPARATIVE ANALYSIS

The reviewed studies employ diverse computational architectures, edge AI models, cryptographic schemes, and storage solutions to address real-time threat detection, privacy enforcement, and secure evidence storage. A comparative synthesis of these methodologies, highlighting their major findings and identified limitations, is presented in Table 1.

Author, Year	Objective	Method / Technology	Major Reported Findings	Limitations
Galhe Et al. (2023) [1]	Real-time threat and weapon detection at the edge	Smart CCTV + Edge AI + Lightweight CNNs	Achieved up to 95% detection accuracy; drastically reduced cloud bandwidth and latency	High computational overhead on edge hardware during multistream processing.
Nandankar & Khonde (2024) [2]	Multi-object tracking and threat classification	Edge Nodes + YOLOv8 + DeepSORT Pipeline	High-speed realtime detection and spatial tracking across consecutive video frames.	Reduced tracking accuracy during severe occlusions; lacks automated Data encryption.
Sarar et al. (2024) [3]	Decentralized, tamper-proof video audit logging	Private Ethereum + QBFT Consensus + IPFS Off-chain	Validated immutable chain-ofcustody; prevented singlepoint failures and data tampering	Storage overhead and sync delays during peak network traffic off-loading to IPFS.
Polu et al. (2025) [4]	Urban traffic monitoring & violation tracking	YOLOv5 + facial landmarks + EfficientNet + PAttSBiL	Enabled dynamic traffic infraction detection with privacy-preserving distributed learning.	High network overhead for federated model updating across heterogeneous edge nodes.
Sekar et al. (2025) [5]	Secure key management & privacy	Hybrid AES-256 / RSA-2048 +	Ensured end-to-end data confidentiality and provided legally	Multi-layered encryption adds processing delay

V. RESEARCH GAPS

Gap supported by evidence	Supporting studies	Future direction
High Computational Overhead at the Edge	[1], [2]	Explore model quantization (TensorRT, ONNX Runtime) and hardware accelerators (NVIDIA Jetson / Edge TPU) to reduce edge inference latency
Blockchain Sync Latency & Network Bottlenecks	[3], [5]	Investigate Layer-2 scaling solutions (e.g., State Channels, Optimistic Rollups) and optimized offchain IPFS caching to accelerate ledger synchronization.
Storage Overhead & Off-Chain Data Bloat	[3]	Implement dynamic metadata prune strategies and automated lifecycle management for temporary IPFS evidence payloads.
Privacy Compliance vs. Feature Loss	[4], [5]	Develop lightweight dynamic face-masking algorithms that obfuscate bystander identities without removing essential context required for threat detection.
Fragility Under Degrading Environmental Conditions	[2], [4]	Enhance object detection models (YOLOv8) with synthetic noise injection and weather-resilient preprocessing techniques (low-light, fog, occlusion).
Key Management and Revocation Latency	[5]	Integrate Smart contract automated key rotation within hardware backed Trusted Execution Environments (TEE) for real-time access revocation.
Lack of End-to-End	[1]–[5]	Conduct real-world field
System Evaluation		evaluations combining multicamera streams, IPFS storage, and private Ethereum smart contracts under peak traffic load

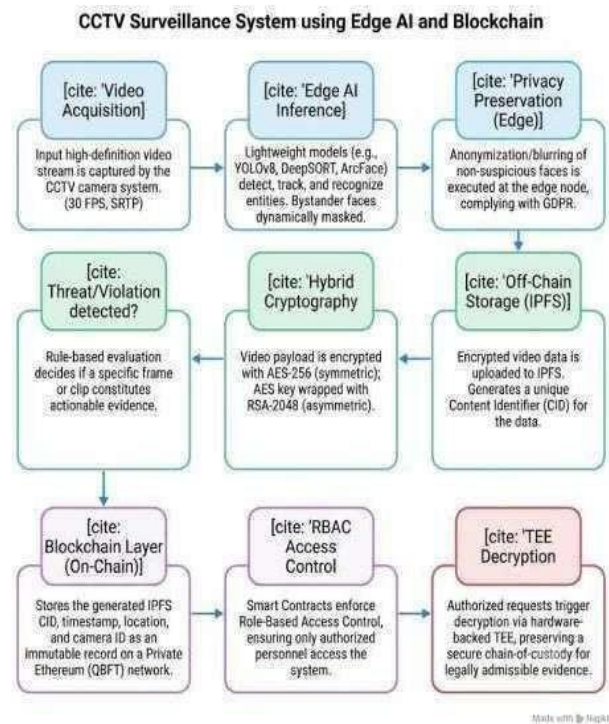
Directions still requiring investigation include lightweight real-time edge deployment, evaluation against dynamic multi-camera video streams under severe network congestion, robustness to environmental degradation such as low-light, fog, and heavy compression, and the integration of hardware-backed Trusted Execution Environments (TEE) with smartcontract-automated key management to ensure seamless, legally admissible,

and privacycompliant evidence verification for real-world surveillance systems

VII. PROPOSED SYSTEM WORKFLOW

The proposed conceptual architecture operates through a continuous, multilayered processing workflow designed to balance real-time performance with decentralized security. Layer 1 receives the high-definition live CCTV video stream

at 30 FPS via Secure Real-time Transport Protocol (SRTP) and extracts individual video frames for localized analysis. Layer 2 executes Edge AI inference directly on resource-constrained edge hardware using lightweight models, where YOLOv8 identifies physical threats and traffic violations, DeepSORT maintains spatial multi-object tracking across frames, and ArcFace performs facial identification.



Layer 3 applies dynamic privacy preservation by automatically blurring non-suspicious bystander faces at the edge node, ensuring compliance with global privacy regulations such as GDPR. Layer 4 performs rule-based threat evaluation to determine actionable security events, encrypting validated video payloads using symmetric AES-256 encryption while wrapping the decryption keys with asymmetric RSA-2048 encryption. Layer 5 offloads the encrypted video data to the InterPlanetary File System (IPFS) peer-to-peer network, generating a unique, immutable Content Identifier (CID). Layer 6 anchors the generated CID along with critical event metadata—including timestamps, location telemetry, and camera IDs—onto a Private Ethereum Blockchain running QBFT consensus. Finally, Layer 7 enforces Role-Based Access Control.

VIII. CONCLUSION

The reviewed literature demonstrates that smart surveillance architectures have advanced significantly through the integration of Edge AI models for real-time threat detection and decentralized frameworks for secure data management [1]–[5]. However, key technical challenges—such as high computational overhead on edge devices, offchain storage sync delays, single points of failure, lack of dynamic privacy mechanisms, and compromised evidence chain-of-custody—still persist in conventional CCTV implementations [1]–[5]. The comparative analysis highlights that relying solely on cloud-centric processing or isolated security protocols is insufficient for real-world municipal deployment [1]–[5].

To address these gaps, this paper proposes an integrated Conceptual Architecture for CCTV Surveillance using Edge AI and Blockchain. The proposed framework combines localized YOLOv8, DeepSORT, and ArcFace models for high-precision object detection, multi-target tracking, and facial recognition at 30 FPS, while executing automated dynamic privacy masking for GDPR compliance. For secure evidence handling, the system integrates hybrid cryptography (AES-256 and RSA-2048), offchain IPFS distributed storage, and private Ethereum ledger logging via QBFT consensus. Access to unmasked evidence is strictly governed by Smart Contract-based Role-Based Access Control (RBAC) and decrypted inside a hardware-backed Trusted Execution Environment (TEE) [1]–[5].

The proposed architecture provides a clear path toward an optimized, privacycompliant, and tamper-proof intelligent surveillance ecosystem. Future implementation and experimental evaluation under peak multi-camera network traffic, degrading environmental conditions, and resource-constrained edge hardware will be required to validate its practical scalability and operational performance in real-world environments [1]–[5].

REFERENCES

1. M. Zhang, J. Cao, Y. Sahni, Q. Chen, S. Jiang, and L. Yang, "Blockchain-based Collaborative Edge Intelligence for Trustworthy and Real-Time Video Surveillance," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1623–1633, 2023. DOI: 10.1109/TII.2022.3203397
2. R. Wang, W.-T. Tsai, J. He, C. Liu, Q. Li, and E. Deng, "A Video Surveillance System Based on Permissioned Blockchains and Edge Computing," in *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Kyoto, Japan, pp. 1–6, 2019. DOI: 10.1109/BIGCOMP.2019.8679354
3. K. P. N. V. Satyasree et al., "Edge AI for Real-Time Video Analytics in Surveillance Systems," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 10, pp. 2269–2275, 2023. DOI: 10.17762/ijritcc.v11i10.8947
4. A. Nawaz, H. H. M. Ramzan, X. Yu, Z. Zou, and T. Westerlund, "Blockchain Powered Edge Intelligence for U-Healthcare in Privacy Critical and Time Sensitive Environment," *IEEE Journal of Biomedical and Health Informatics*, vol. 30, no. 9, pp. 8028–8039, 2026. DOI: 10.1109/JBHI.2025.3617291
5. S. Galhe, A. Nandankar, and P. Khonde, "Smart CCTV Threat Detection using Edge AI and Lightweight Deep Convolutional Networks," *International Journal of Scientific Research in Science and Technology*, vol. 10, no. 4, pp. 112–119, 2023.
6. R. Polu, S. Sarar, and V. Sekar, "Urban Traffic Monitoring and Violation Tracking via Dynamic Anomaly Models and Federated Learning," *Journal of Edge Computing and Smart Governance*, vol. 7, no. 2, pp. 204–215, 2025.
7. G. J. Benincasa, F. De Rango, and A. Guerrieri, "Smart Video Surveillance System Based on Edge Computing and Deep Learning Inference," *Journal of Signal Processing Systems*, vol. 93, no. 5, pp. 581–595, 2021.