INTERNATIONAL MULTI DISCIPLINARY JOURNAL FOR RESEARCH & DEVELOPMENT
eISSN 2394-6334 Impact factor: 8.854 Volume 13, issue 04 (2026)

# METHODS OF ANALYZING BIG DATA FROM VIDEO SURVEILLANCE SYSTEMS IN SMART CITIES

Author: PhD **Tursunbek Sadriddinovich Jalolov**

Institution: Asia International University

E-mail: tursunbekjalolov@gmail.com

**Abstract.**

This article examines methods of analyzing large-scale data (Big Data) obtained from video surveillance systems in smart cities. Technologies for data collection, storage and processing, deep learning algorithms, and real-time analysis systems are reviewed. Practical application areas and existing challenges are analyzed.

***Keywords:*** *Big Data, video surveillance, smart city, deep learning, data analysis, cloud technologies, artificial intelligence, real-time analysis.*

## Introduction

The sharp increase in the number of video surveillance cameras in modern cities results in terabytes of data being generated daily. According to international studies, a single IP camera generates an average of 15–40 GB of video data per day. In major cities with thousands of cameras installed, this figure reaches petabytes. Processing and analyzing such massive volumes of data using traditional methods is impossible, which has given rise to the need for Big Data technologies.

Big Data refers to datasets whose volume, velocity, and variety exceed the capabilities of traditional databases. In the context of video surveillance, Big Data encompasses not only video recordings but also metadata (time, location, camera identifier), lists of objects detected by artificial intelligence, movement trajectories, and anomaly signals.

The purpose of this article is to scientifically examine modern methods, technologies, and algorithms for analyzing the large volumes of data generated by video surveillance systems in smart cities. The research employs systematic literature analysis, comparative, and classification methods.

## The concept of big data and characteristics of video surveillance data

The concept of Big Data is typically characterized through the "5V" model: Volume, Velocity, Variety, Veracity, and Value. Video surveillance data fully possesses all five characteristics, making it one of the most complex objects of Big Data analysis.

***Table 1. The "5V" Model of Big Data in the Context of Video Surveillance***

| ***Characteristic*** | ***Description*** | ***Video Surveillance Example*** |
|---|---|---|
| ***Volume*** | ***Extremely large data volumes*** | ***One camera: 15–40 GB/day; petabytes at city scale*** |

| | | |
|---|---|---|
| *Velocity* | *High-speed data flow* | *Real-time 25–30 fps; thousands of parallel streams* |
| *Variety* | *Diverse data formats* | *Video, images, metadata, GPS coordinates, sensors* |
| *Veracity* | *Data quality and accuracy* | *Quality depends on lighting, weather, camera angle* |
| *Value* | *Extracting value from data* | *Crime detection, traffic optimization, forecasting* |

A distinctive characteristic of video surveillance data is that it is predominantly unstructured (video, images). Traditional relational databases are unable to efficiently process such data. Therefore, modern solutions such as NoSQL databases (MongoDB, Cassandra), distributed file systems (HDFS), and object storage systems (Amazon S3, MinIO) are employed.

Another important aspect of video surveillance Big Data is the temporal dependency of the data. Each frame in a video stream is semantically linked to the preceding and following frames. This characteristic requires analysis algorithms to account for not only the spatial but also the temporal dimension.

**Data collection and storage technologies**

To effectively analyze Big Data obtained from video surveillance systems, a proper infrastructure for data collection, transmission, and storage is essential. In modern smart city architecture, data flow is organized across three main layers: the edge layer, the network layer, and the cloud layer.

At the edge layer, preliminary processing is performed directly on the cameras — video compression (H.264, H.265, H.266 codecs), motion detection, and initial object classification. At the network layer, data is transmitted to central servers via 5G, fiber-optic, or Wi-Fi 6 communication channels. At the cloud layer, data is stored long-term and subjected to in-depth analysis.

*Table 2. Comparison of Video Surveillance Data Storage Technologies*

| *Technology* | *Type* | *Advantages* | *Disadvantages* |
|---|---|---|---|
| *HDFS (Hadoop)* | *Distributed file system* | *Large capacity, reliability, scalability* | *Slow for real-time* |

| | | | |
|---|---|---|---|
| *Apache Cassandra* | *NoSQL database* | *High write speed, fault tolerance* | *Complex queries limited* |
| *Amazon S3 / MinIO* | *Object storage* | *Unlimited scalability, low cost* | *High latency* |
| *Redis / Memcached* | *In-memory cache* | *Highest speed, ideal for real-time* | *Capacity limited by RAM* |
| *Elasticsearch* | *Search engine* | *Fast metadata search, indexing* | *Not suitable for video storage* |

In practice, a hybrid approach is often employed: high-speed caches like Redis for real-time analysis, Elasticsearch for metadata search, and HDFS or object storage systems for video archiving. This architecture is based on design patterns known as Lambda Architecture or Kappa Architecture.

Lambda Architecture consists of two parallel layers: the batch layer performs deep analysis of historical data, while the speed layer processes real-time streams. Kappa Architecture, in contrast, uses a single stream processing layer, making it simpler to implement and maintain. In video surveillance systems, Lambda Architecture is often preferred, as the need for retrospective analysis of historical video archives frequently arises.

**Analysis methods and algorithms**

Methods for analyzing video surveillance Big Data can be divided into three main groups: traditional computer vision methods, deep learning-based methods, and hybrid approaches. Each group has its own advantages and application areas.

Traditional computer vision methods include background subtraction, optical flow, and feature extraction (HOG, SIFT, SURF). These methods require fewer computational resources but suffer from reduced accuracy in complex conditions (lighting changes, occlusions).

Deep learning-based methods have revolutionized video surveillance analysis in recent years. Convolutional neural networks (CNNs) automatically learn features from images. For object detection, YOLO (You Only Look Once), SSD (Single Shot Detector), and Faster R-CNN are the most widely used algorithms. For video analysis, 3D-CNNs, SlowFast Networks, and Video Transformers are employed.

***Table 3. Comparison of Key Video Surveillance Analysis Algorithms***

| *Algorithm* | *Task* | *Speed (FPS)* | *Accuracy (mAP)* | *Features* |
|---|---|---|---|---|
| *YOLOv8* | *Object detection* | *45–80* | *53.9%* | *Real-time, lightweight architecture* |
| *Faster R-CNN* | *Object detection* | *5–15* | *42.7%* | *High accuracy, slow* |
| *DeepSORT* | *Object tracking* | *30–40* | — | *Tracking + re-identification* |
| *3D-CNN (I3D)* | *Action recognition* | *10–20* | *74.2%** | *Temporal analysis, resource-intensive* |
| *Video Swin Transformer* | *Video understanding* | *8–15* | *84.6%** | *Highest accuracy, heavy* |

**Results measured on the Kinetics-400 dataset*

Hybrid approaches combine the advantages of traditional and deep learning methods. For example, moving objects are first isolated through background subtraction, then classified using CNNs. This approach conserves computational resources while maintaining high accuracy.

Distributed computing frameworks such as MapReduce, Apache Spark, and Apache Flink enable parallel processing of large volumes of video data. Apache Spark Streaming and Flink support real-time stream processing, which is essential for simultaneously processing video streams from thousands of cameras.

The Edge AI approach holds a special place in video surveillance Big Data analysis. Specialized chips such as NVIDIA Jetson, Intel Movidius, and Google Coral can be installed in cameras, enabling real-time object detection, facial recognition, and anomaly detection algorithms. This approach reduces the volume of data transmitted to central servers by 70–90% and eliminates network latency.

**Practical application areas**

Video surveillance Big Data analysis is yielding practical results across various domains of the smart city. In public safety, predictive policing systems analyze historical video surveillance data to identify areas with a high probability of crime. The cities of Los Angeles and Chicago have tested these systems and reported a 15–20% decrease in crime rates in certain areas.

In transportation, video surveillance Big Data is used for traffic flow optimization, congestion forecasting, and automatic detection of road accidents. Singapore has successfully implemented an adaptive traffic light management system by analyzing data from thousands of cameras in real time, resulting in a 12% reduction in average travel time.

In retail, video surveillance analysis is used for customer movement heat mapping, identifying visitation trends, and theft prevention. Amazon Go stores have implemented cashier-less shopping through a combination of computer vision and sensors — one of the most advanced practical applications of video surveillance Big Data.

In healthcare, hospital video surveillance data is analyzed for patient safety monitoring, fall detection, and automated verification of sanitation compliance. During the COVID-19 pandemic, thermal cameras and facial recognition systems were widely used for remote temperature measurement and mask-wearing verification.

In Uzbekistan, video surveillance Big Data analysis is being practically applied within the "Safe City" project. Data from thousands of cameras installed in Tashkent is collected at a central situation center and analyzed in real time. This system performs automatic detection of traffic violations, vehicle search, and public order monitoring.

In urban planning, video surveillance Big Data is used to analyze the temporal and spatial distribution of pedestrian and traffic flows, identify the need for new roads and bridges, and evaluate the efficiency of existing infrastructure. The city of Barcelona has integrated video surveillance data with GIS (Geographical Information Systems) for use in the urban planning process.

In emergency management, video surveillance Big Data is critically important for planning population evacuations during natural disasters, fires, and industrial accidents, coordinating rescue operations, and assessing situations in real time. In Tokyo, data from video surveillance cameras is automatically analyzed during earthquakes, with damaged buildings and areas identified within minutes.

**Challenges and solutions**

A number of serious challenges exist in the analysis of video surveillance Big Data. These challenges are of a technical, legal, and social nature, and appropriate solutions are being developed for each.

***Table 4. Key Challenges and Solutions in Video Surveillance Big Data Analysis***

| ***Problem*** | ***Description*** | ***Proposed Solutions*** |
|---|---|---|
| ***Storage volume*** | ***Storing petabytes of video*** | ***Smart compression*** |

| | | |
|---|---|---|
| | *data is expensive and complex* | *(H.266/VVC), time-based deletion policies, tiered storage* |
| *Real-time processing* | *Analyzing streams from thousands of cameras simultaneously* | *Edge computing, GPU clusters, Apache Flink stream processing* |
| *Privacy* | *Risk of intrusion into citizens' private lives* | *Data anonymization, federated learning, GDPR-compliant architecture* |
| *Algorithmic bias* | *AI models performing differently across demographic groups* | *Diversified training datasets, fairness metrics, regular audits* |
| *Cybersecurity* | *Risk of intercepting or manipulating video streams* | *End-to-end encryption, blockchain-based authentication* |
| *Energy consumption* | *GPU-based analysis systems require significant energy* | *Energy-efficient chips (NPU), model pruning, quantization* |

Federated learning technology offers an innovative solution to the privacy challenge. In this approach, the AI model is trained locally on each camera or edge device rather than on a central server. Only model parameters (gradients) are sent to the center, while the original video data never leaves the device. Google and NVIDIA are adapting this technology for the video surveillance domain.

Model compression techniques — pruning, quantization, and knowledge distillation — enable large neural networks to run on edge devices. For example, the YOLOv8-nano model is 4 times smaller than the full YOLOv8, yet accuracy decreases by only 3–5%. This makes real-time analysis on edge computing devices practically feasible.

**Conclusion**

Based on the research findings, the following conclusions can be drawn. Video surveillance systems in smart cities are among the largest sources of modern Big Data, and their

effective analysis can fundamentally improve the quality of urban governance. The "5V" model encompasses all characteristics of video surveillance data and serves as a key benchmark in designing analysis systems.

A hybrid architecture (edge + cloud) represents the most effective approach for data collection and storage. The combination of HDFS, Cassandra, and object storage systems satisfies diverse analytical needs. Among analysis algorithms, deep learning methods — particularly YOLOv8 and Video Transformer models — demonstrate the highest performance in both accuracy and speed.

However, challenges such as privacy, cybersecurity, algorithmic bias, and energy consumption must be addressed. Innovative technologies including federated learning, model compression, and blockchain offer effective solutions to these problems. In Uzbekistan, the "Safe City" project serves as an important platform for the practical implementation of these technologies.

Future research should continue with empirical data from specific urban infrastructure cases, as well as experimental investigation of the impact of federated learning and edge AI approaches on video surveillance analysis efficiency.

Additionally, the development of local AI models for video surveillance Big Data analysis in Uzbekistan, expanding the capabilities of national data centers, and training specialists in this field require special attention. Supporting scientific research in this direction within the Digital Uzbekistan strategy will be a significant factor in enhancing the country's competitiveness.

**References.**

1. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. Communications of the ACM, 51(1), 107–113.

2. Redmon, J., et al. (2016). You Only Look Once: Unified, Real-Time Object Detection. CVPR, 779–788.

3. Ren, S., et al. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. NeurIPS, 91–99.

4. Zaharia, M., et al. (2016). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56–65.

5. McMahan, B., et al. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS, 1273–1282.

6. Liu, Z., et al. (2022). Video Swin Transformer. CVPR, 3202–3211.

7. Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? A new model and the Kinetics dataset. CVPR, 6299–6308.

8. Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. ICIP, 3645–3649.

9. Hashem, I. A. T., et al. (2016). The role of big data in smart city. International Journal of Information Management, 36(5), 748–758.

10. Zanella, A., et al. (2014). Internet of things for smart cities. IEEE Internet of Things Journal, 1(1), 22–32.

11. Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1–14.

12. Elmaghraby, A. S., & Losavio, M. M. (2014). Cyber security challenges in smart cities. Journal of Advanced Research, 5(4), 491–497.

13. NIST. (2019). Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects. NIST IR 8280.

14. Gartner, Inc. (2023). Market Guide for Video Surveillance.

15. IHS Markit. (2024). Global Video Surveillance Market Report.

16. Decree of the President of the Republic of Uzbekistan on the "Digital Uzbekistan — 2030" Strategy. (2020).

17. Vaswani, A., et al. (2017). Attention Is All You Need. NeurIPS, 5998–6008.

18. Lin, T.-Y., et al. (2014). Microsoft COCO: Common objects in context. ECCV, 740–755.