Modern Management Science & Engineering
ISSN ISSN 2052-2576
Vol. 8, No. 3, 2026
www.scholink.org/ojs/index.php/mmse

Original Paper

Research on All-Domain Situation Awareness Technology for Video Surveillance Based on Artificial Intelligence

Zheyuan Wang[1*], Buwei Zhang[2], & Linjun Li[3]

[1] Nanjing Research Institute of Electeonic, No. Lingshan North Road, Qixia District, Nanjing, China

[2] Beijing-based Training Center, China

[3] Nanjing Research Institute of Electeonic, China

[*] Corresponding Author: harrywzy@foxmail.com

Received: May 19, 2026 Accepted: June 22, 2026 Online Published: July 7, 2026

doi:10.22158/mmse.v8n3p93 URL: http://dx.doi.org/10.22158/mmse.v8n3p93

Abstract

In the context of the continuous deepening of smart city and safe city construction, the urban video surveillance system has achieved all-round three-dimensional coverage, and the traditional monitoring operation mode relying on manual duty and post-event review has become difficult to adapt to the current dynamic and complex public security control requirements. As a core research direction in the field of intelligent security, global situation awareness, relying on computer vision, artificial intelligence and spatio-temporal big data analysis technology, breaks through the functional limitations of traditional video surveillance and can conduct intelligent perception, analysis and early warning of various scene elements in the monitored area at all times and in all directions, It effectively addresses industry pain points such as fragmented vision, low data utilization, lagging event response, and frequent false and missed detections in traditional surveillance. This paper, in light of the current development status of intelligent surveillance technology, systematically explores the integrated application system of artificial intelligence and video global situation awareness, comparatively analyzes the inherent flaws of traditional surveillance models and the technical advantages of AI intelligent perception, and focuses on dissecting core key technologies such as lightweight target detection, multi-target trajectory tracking, abnormal behavior recognition, and multi-source data fusion. Build a five-layer modular global situation awareness system architecture. By setting up a multi-scenario experimental environment for comparative testing, verify the practicality and accuracy of the technical solution proposed in this paper, sort out the shortcomings and bottlenecks of the current technology in combination with the current application status, propose targeted optimization and improvement plans, and predict the future development trend of the industry. The experimental and *research results show that artificial intelligence technology can effectively enhance the all-domain scene perception ability of video surveillance, the accuracy of complex event judgment and the efficiency of emergency response, provide technical support for the intelligent upgrade of scenarios such as urban traffic governance, public security prevention and control, and emergency management and control of sudden incidents, and provide theoretical reference for the standardized and large-scale construction of the all-domain intelligent security system.*

Keywords

Artificial intelligence, Video surveillance, Global situational awareness, Computer vision, Spatio-temporal big data, Smart security

1. Introduction

In recent years, China has made steady progress in new urbanization and fine urban governance. Various regions have continuously advanced the layout of high-definition and ultra-high-definition video surveillance points and gradually built three-dimensional video surveillance networks covering key areas such as main and secondary roads, commercial districts, industrial parks, transportation hubs, and residential communities. Video surveillance equipment has become the core infrastructure for urban public security governance and daily order control. The deployment of a vast number of video capture devices has accumulated a huge amount of basic data resources for urban scene situation analysis, but it has also given rise to significant problems of "data redundancy, lack of information, and lagging judgment". The core functions of traditional video surveillance systems are limited to video capture, local storage and manual review and verification. The overall operation is highly dependent on human presence and can only achieve post-event traceability of emergencies, lacking the ability to predict in advance and control during the event. In the face of large-scale, multi-target, and highly dynamic complex surveillance scenarios, traditional surveillance systems have many problems such as fragmented vision, incomplete perception of scene elements, limited accuracy in identifying abnormal events, and inability to link and analyze cross-point data, making it difficult to meet the core demands of modern public security governance for proactive prevention and control, overall planning, and rapid response.

2. Architecture Design of Video global Situation Awareness System Based on AI

Combining the functional requirements of global situation awareness, the characteristics of application scenarios and the features of artificial intelligence technology, this paper designs a hierarchical and modular architecture of intelligent video surveillance global situation awareness system. In the existing video intelligent perception algorithm system, traditional recognition algorithms are mainly divided into two categories. One category is the two-stage detection algorithm represented by SSD and Faster R-CNN. This type of algorithm has high detection accuracy, but has a large number of network parameters, time-consuming reasoning, strict requirements for terminal computing power, and is difficult to adapt to real-time perception scenarios of edge devices; The other is the early single-stage algorithms represented by YOLOv3 and YOLOv4. Although they have a faster inference speed, they have a high missed detection rate and weak generalization ability for common small targets, occluded targets, and dense crowd targets in surveillance scenarios, and are not robust enough in complex outdoor surveillance scenarios. It fails to meet the core requirements of high precision, low latency and strong adaptability for global situational awareness. Compared with traditional algorithms, the YOLOv8 algorithm has significant advantages in network lightweighting, feature extraction capabilities, and multi-scale target detection performance. It can balance detection accuracy and inference speed, and is more suitable for the perception requirements of multi-scenario, multi-target, and high real-time performance in all-domain video surveillance. This is also an important basis for the system in this paper to choose this algorithm as the core perception model. (Zhang, Wang, & Li, 2023)

Based on the algorithmic characteristics and scenario requirements mentioned above, this paper builds a top-down five-layer architecture, namely the data collection layer, edge preprocessing layer, AI intelligent perception layer, situation assessment layer, and application service layer. Each layer has independent functions, logical connections, and collaborative operations, forming a complete closed-loop system of "data collection - preprocessing - intelligent perception - situation assessment - application output". To further illustrate the operational logic and practical value of each level, this paper combines two typical scenarios, namely urban road traffic and full-domain security in commercial plazas, and incorporates specific application cases to support the rationality and practicality of the architecture design. In the urban morning rush hour traffic control scenario, the system continuously collects video data of vehicle and pedestrian flow through all-domain road monitoring equipment, completes image noise reduction and invalid frame elimination through edge devices, detects vehicle, pedestrian and non-motorized vehicle targets in real time relying on the YOLOv8 model, statistically analyzes the traffic density of road sections in combination with the trajectory tracking algorithm, judges the congestion situation and pushes early warning information; In the security scenarios of large commercial plazas, the system covers key locations such as plaza entrances and exits, escalator passageways, and open public areas. It monitors abnormal behaviors such as people gathering, lingering, running and chasing in real time, and links the background to complete risk classification assessment and device-linked early warning to achieve dynamic control of the entire area. The architecture has a clear hierarchy and strong scalability, and can adapt to various application scenarios such as urban security, traffic control, park management, and emergency response. The overall architecture of the system is shown in Figure 1.

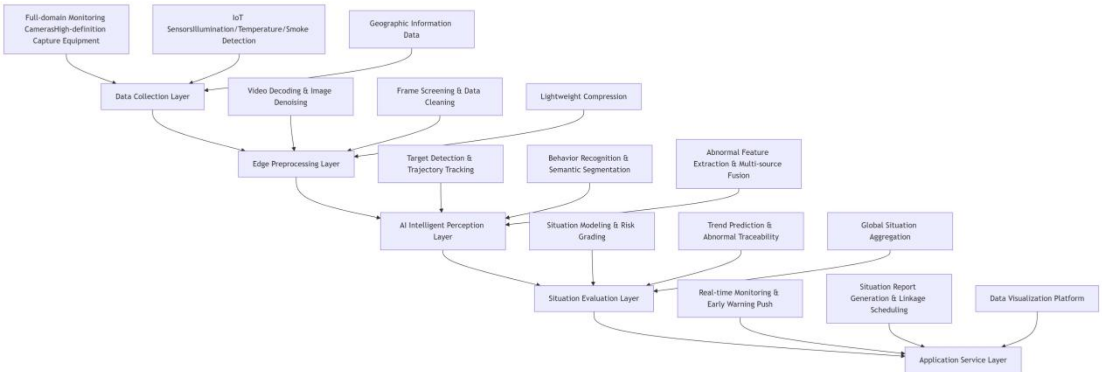

Figure 1. Architecture Diagram of the AI-based Video Surveillance Global Situation Awareness System

2.1 Data Acquisition Layer

The data acquisition layer is the data foundation of the global situation awareness system and is mainly responsible for the real-time collection of multi-dimensional basic data of the global scenarios. The data sources collected by the system are mainly divided into two categories. The first category is core video image data, which relies on the all-domain layout of high-definition and ultra-high-definition network cameras, smart dome cameras, and road surface capture devices to collect real-time video streams and static scene images from various areas from multiple angles and without interruption; The second category consists of auxiliary perception data, including scene geographic information, environmental sensor data, equipment operation status data, etc., providing multi-dimensional data support for global situation modeling. This level ensures the integrity, real-time nature and comprehensiveness of subsequent situation analysis data through the all-day, blind-spot-free data collection mode. (Liu & Chen, 2023)

2.2 Edge Preprocessing Layer

The direct upload of massive high-definition video streams to the cloud platform can easily cause problems such as network bandwidth congestion, cloud computing power overload, and excessive data transmission delay. To address these issues, the system adds an edge preprocessing layer to perform data lightweight processing in front-end edge computing devices. The main functions of this layer include video stream decoding, picture noise reduction, adaptive light and shadow correction, invalid frame filtering, redundant data cleaning and video compression, eliminating blurry, severely occluded and invalid repetitive redundant data, and retaining core scene feature information. At the same time, it completes the initial alignment and standardization processing of heterogeneous data from multiple devices, providing high-quality data support for the upper-level AI intelligent analysis module, effectively reducing cloud computing power consumption, and significantly improving the real-time performance and stability of the overall system operation. (Wang & Zhou, 2023)

2.3 AI Intelligent Perception Layer

The AI intelligent perception layer is the core technology of the entire system, relying on deep learning machine vision algorithms to complete the intelligent perception and analysis of all elements of the video scene. This layer consists of six core functional modules, namely object detection, multi-object trajectory tracking, human behavior recognition, scene semantic segmentation, abnormal feature extraction, and multi-source data fusion. The system uses a lightweight improved YOLO deep learning model to achieve precise positioning and classification detection of targets such as people, vehicles, clutter, and dangerous goods within the scene; Combined with the DeepSORT algorithm for multi-target cross-frame continuous trajectory tracking, accurately reproducing the dynamic motion of all targets in the domain; Based on the spatio-temporal feature fusion algorithm, identify various abnormal behaviors within the scene; By using semantic segmentation technology to divide scene units such as roads, crowds, buildings, and alert areas, fine-grained scene perception is achieved. (Li & Zhao, 2024)

2.4 Situation Assessment Layer

The situation analysis layer has achieved a functional upgrade of the system from "element perception" to "global situation cognition", and it is the core decision-making level of the all-domain situation awareness system. Based on the structured data output by the AI intelligent perception layer, this level builds an all-domain scenario situation analysis model and focuses on four core tasks: situation aggregation, risk classification, anomaly traceability, and trend prediction. The system integrates perception data from all monitoring points across the domain, summarizes and generates regional global situation maps, and combines preset scene control rules to complete the risk classification of various abnormal events into three levels: general, early warning, and high risk. At the same time, it uses spatio-temporal big data to mine the changing patterns of the scene situation, predict the development trend of regional risks, and provide scientific decision-making basis for subsequent precise emergency response and regular control.

2.5 Application Service Layer

The application service layer is the functional output port of the system to end users, mainly undertaking functions such as situation data visualization, risk warning push, device linkage handling and data operation and maintenance management. Specifically, it includes real-time visualization display of the entire domain situation, pop-up warning of abnormal events, message push between mobile terminals and background terminals, emergency equipment linkage dispatch, automatic generation of situation statistics reports, historical situation backtracking query, equipment status operation and maintenance management, etc. Fully adapted to the daily duty of control personnel, emergency response to unexpected incidents, situation data analysis, equipment operation and maintenance control and other work requirements, to form a full-process intelligent work closed loop of perception, judgment, early warning, handling and review. (Chen & Wu, 2023)

3. Research on Core Artificial Intelligence Technologies for Global Situational Awareness

3.1 Lightweight Target Detection Technology

Object detection is a fundamental prerequisite for achieving global situational awareness, and its core role is to complete the positioning, classification, and feature extraction of various types of targets in surveillance footage. Traditional deep learning detection models have complex network structures, consume a lot of computing power, and have slow inference speeds, making it difficult to adapt to the real-time detection requirements of edge terminal devices and unable to support the continuous perception of dynamic scenes across the entire domain. In this paper, the YOLOv8 lightweight model, which has undergone structural simplification and parameter optimization, is selected. By eliminating redundant convolutional layers, introducing channel attention mechanisms, and optimizing the model loss function, the inference operation speed of the model is significantly improved while basically ensuring the accuracy of target detection. The optimized model can adaptively adapt to various complex picture quality scenarios such as high definition, normal picture quality, low light, and weak exposure, effectively improving the problems of missed and false detections of dense targets, small targets, and occluded targets, achieving real-time and accurate detection of multiple targets in the entire domain, and providing stable and reliable basic data for subsequent trajectory tracking, behavior analysis, and situation judgment. (Qaraqe, Elzein, & Basaran, 2024)

3.2 Multi-target Spatiotemporal Trajectory Tracking Technology

The core of global situation awareness is to grasp the dynamic change patterns of all moving targets within the area. Relying solely on static target detection in a single frame cannot reflect the dynamic evolution of the scene. Therefore, dynamic awareness must be achieved by relying on multi-target spatio-temporal trajectory tracking technology. This paper uses a fusion scheme of the YOLOv8 detection model combined with the DeepSORT trajectory matching algorithm to extract target-specific features through the detection model and complete cross-frame target association matching using the trajectory matching algorithm to continuously record dynamic information such as the real-time position, movement speed, and travel trajectory of each target. This algorithm can effectively solve the problem of trajectory interruption and identity confusion caused by target overlap and screen occlusion, and accurately distinguish different moving targets. By integrating trajectory data from multiple points across the entire area, it can fully restore the dynamic situation of the movement of people and the passage of vehicles in the entire monitored area, achieving precise perception of the dynamic information of the entire scene. (Zhao & Sun, 2023)

3.3 Intelligent recognition technology for Abnormal Behaviors in Complex Scenarios

Abnormal behavior recognition is a key link in risk situation awareness, and its recognition accuracy directly determines the effectiveness and reliability of system early warning. Traditional behavior recognition algorithms can only identify regularized and simplified fixed behavior patterns, and have poor ability to identify complex dynamic scenarios and latent abnormal behaviors, with insufficient adaptability. This paper uses a spatio-temporal feature fusion deep learning framework, relying on a 3D convolutional neural network to synchronously extract the spatial structure features and time series dynamic features of video frames, and combines a self-built complex scene behavior dataset to complete model training and iterative optimization. The optimized model can accurately identify a variety of typical abnormal events such as people gathering, fast running, climbing over barriers, brawling and pulling, illegal retention, vehicle driving in the wrong direction, people falling to the ground, fire smoke, clutter accumulation, etc. At the same time, it supports customizing abnormal identification rules according to the control requirements of different scenarios and can flexibly adapt to various differentiated control scenarios. (Lin, 2024)

3.4 Multi-source Data fusion Global Situation Modeling Techniques

Single video data has the drawbacks of a single perception dimension and weak environmental anti-interference ability. Relying solely on video images cannot fully restore the real situation of the scene, and problems such as one-sided situation judgment and risk misjudgment are prone to occur. To address this issue, this paper employs multi-source data fusion technology, integrating all-domain video perception data, environmental sensing data, geographic information data, and historical situation statistics data, and completes the cleaning, alignment, and fusion processing of multi-source heterogeneous data through weighted fusion algorithms to eliminate data redundancy and data conflict, and builds a unified situation data model for the entire domain. Based on this model, dynamic statistics and visual presentation of situation information such as regional pedestrian flow, vehicle flow, risk point distribution, and frequency of abnormal events can be completed. At the same time, combined with spatio-temporal big data time series analysis methods, the evolution trend of scene situation can be deduced, truly achieving the upgrade and transformation of monitoring mode from passive post-event traceability to active pre-event prediction.

3.5 Real-time Intelligent Reasoning Technology at the Edge

The layout of video surveillance points is dense and the volume of real-time video stream data is huge. If the cloud-based centralized reasoning mode is adopted, problems such as network bandwidth congestion, excessive reasoning latency, video picture stammering and disconnection are very likely to occur, which cannot meet the real-time requirements of global situation awareness. This paper uses technical means such as model lightweight clipping, fixed-point quantification compression, and operator adaptation optimization to lightweight transform the AI inference model and directly deploy the optimized model on front-end edge computing devices. It enables real-time local parsing of video data, local anomaly identification and early warning, and uploads only core situation data and anomaly event data to the cloud platform for aggregated analysis, significantly reducing network bandwidth resource consumption. The inference mode, which has been tested and optimized, can stably control the overall delay within 100ms, fully guaranteeing the real-time performance and operational stability of global situation awareness. (Huang & Zheng, 2023)

4. Technical Performance Experiments and Results Analysis

To verify the feasibility, accuracy and real-time performance of the AI video all-domain situation awareness technology solution designed in this paper, a standardized experimental test environment was established in this paper, and three typical monitoring scenarios with high frequency and complexity, namely urban main roads, commercial plazas and entrances and exits of industrial parks, were selected for comparative experiments. A multi-dimensional performance comparison was conducted between the optimized scheme of this paper and the traditional manual monitoring scheme and the basic AI monitoring scheme. Four core indicators, namely target detection accuracy, abnormal event recognition accuracy, system reasoning delay, and false alarm rate, were selected as evaluation criteria to comprehensively test the application effect of the technology. (Tian & Li, 2024)

4.1 Experimental Environment and Test Protocol

Hardware configuration for this experiment: The edge computing device is equipped with an ARM architecture processor, 8G of running memory, and 16 high-definition surveillance cameras are deployed to fully cover the test experiment area for blind-spot-free monitoring. The software runtime environment is based on Python3.9 programming language, PyTorch deep learning framework and OpenCV video processing library. The experimental dataset uses a self-built complex scene video dataset, including 1000 video clips under different lighting conditions, different pedestrian and vehicle flow densities, and different weather conditions, covering various scenarios such as regular passage, static duty, and various abnormal events, ensuring the comprehensiveness of the test data. Three control schemes were set up in the experiment, namely Scheme 1 traditional manual monitoring and duty mode, Scheme 2 basic AI intelligent monitoring mode, and Scheme 3 the all-domain situation awareness AI technology scheme optimized in this paper, and performance indicators were statistically analyzed through multiple sets of repeated experiments.

4.2 Comparative Analysis of Experimental Data

Through multi-scenario repeated testing, data statistics and mean calculation, the comparison results of core performance indicators of the three experimental schemes were obtained. The specific data are shown in Table 1 below.

Table 1. Performance Comparison of Different Video Surveillance Technology Schemes

Test Scheme	Target detection accuracy (%)	Anomaly recognition accuracy (%)	System inference delay (ms)	False positive rate (%)
Traditional manual monitoring schemes	82.3	75.6	Manual delay ≥3000	18.2

Basic AI monitoring solutions	91.5	86.2	320	7.5
This article on Global situational awareness AI solutions	97.8	95.3	95	2.1

In combination with the multiple sets of experimental statistics in Table 1, it is clear that there are significant hierarchical differences in core performance indicators between traditional manual monitoring, basic AI monitoring and the all-domain situation awareness AI solution optimized in this paper. The solution in this paper has achieved all-round breakthroughs in detection accuracy, recognition accuracy, operational real-time performance and early warning stability. It effectively compensates for the many deficiencies of the first two traditional schemes. In the dimension of target detection, traditional manual monitoring is affected by human concentration and visual limitations, with a detection accuracy of only 82.3%, and is prone to visual omissions when dealing with small-sized targets, occluded targets, and densely arranged targets; The basic AI monitoring solution, relying on traditional deep learning algorithms, has improved detection accuracy to 91.5%, but the algorithm is not lightweight enough and has poor adaptability to multi-scale targets in complex scenes, still having some problems of missed detections and false detections. In this paper, based on the YOLOV8-optimized global situation awareness solution, by enhancing multi-scale feature extraction and optimizing the network structure, the target detection accuracy is improved to 97.8%, significantly enhancing the recognition integrity and accuracy of various targets in complex monitoring scenarios.

In the dimension of abnormal event identification, the manual monitoring mode is highly dependent on the experience and concentration of the on-duty personnel. Long-term on-duty is prone to fatigue and burnout. The accuracy rate of abnormal event identification is only 75.6%, and its ability to identify hidden and subtle violations and abnormal behaviors is extremely poor, which cannot meet the needs of refined security control. The basic AI monitoring solution can only identify explicit abnormal behaviors with distinct features. It has limited accuracy in identifying compound and progressive abnormal behaviors in complex scenarios, with an accuracy rate of only 86.2%, and the algorithm has weak generalization ability and insufficient scenario adaptability. This paper integrates spatio-temporal feature fusion algorithms with self-built complex scene datasets for iterative optimization, achieving an abnormal event recognition accuracy of 95.3%. It can not only accurately identify conventional explicit anomalies, but also effectively capture implicit risk behaviors such as slow gathering of people and abnormal retention, meeting the risk identification requirements of various complex monitoring scenarios. (WANG, ZHANG, ZHANG et al., 2022)

In terms of system real-time performance and early warning stability, manual monitoring has a significant response lag. The entire process from picture observation, anomaly verification to early warning response takes no less than 3000ms. It is completely impossible to achieve real-time control and can only be used for post-event traceability. The basic AI monitoring solution relies on the cloud-based inference mode. Affected by network bandwidth and computing power latency, the inference latency is stable at around 320ms. In high-concurrency, multi-target dense scenarios, problems such as latency fluctuations and picture stutters are prone to occur, making it difficult to support full-domain dynamic real-time perception. This paper uses a lightweight edge-end inference architecture to stably control the overall inference latency within 100ms, and the millisecond-level inference speed enables real-time updates of the scene situation and second-level alerts for anomalies. Meanwhile, the system's false alarm rate is controlled at 2.1%. Compared with the high false alarm rate of 18.2% for manual monitoring and 7.5% for basic AI monitoring, it significantly filters out invalid early warning information, avoids problems such as waste of control resources and staff early warning fatigue caused by high-frequency false alarms, and significantly improves the effectiveness and practicality of intelligent security control.

5. Technical Application Difficulties and Optimization Strategies

5.1 Current Technical Application Difficulties

Based on the results of this experimental test and the actual application research in the industry, although AI-based video surveillance all-domain situation awareness technology has achieved initial large-scale application at present, there are still many technical difficulties in complex scene adaptation, data fusion, risk prediction, security protection, computing power support, etc. First, the anti-interference ability in complex environments is insufficient. In harsh scenarios such as rain, snow, fog, direct strong light, low light at night, and picture occlusion, the accuracy of target detection and anomaly recognition of the model will decline significantly, easily causing problems of missed alarms and false alarms, affecting the accuracy of situation analysis. Second, it is difficult to achieve all-domain collaborative integration of multiple devices. There are significant differences in brands, models, and parameters of surveillance devices on the market, and the data formats and transmission protocols output by various devices are not uniform, resulting in obvious data silos and low efficiency of all-domain multi-source data fusion, making it impossible to achieve all-domain coordinated analysis of the situation. Third, the ability to identify latent anomalies is weak. For low-probability, latent anomalies such as slow gathering of people, long-term anomalies lingering, and progressive evolution of risks, the model's feature extraction ability is insufficient, and the accuracy of risk prediction is limited, making it difficult to achieve early warning in advance. Fourth, there are prominent risks to data security and privacy. The all-domain video data contains a large amount of private data such as public facial information and behavioral trajectories, and there are security risks of data leakage and misuse throughout the entire process of data collection, transmission, storage and analysis. Fifth, edge computing power resources are limited. In scenarios of large-scale all-domain monitoring and high-concurrency video stream processing, edge devices have insufficient computing power reserves, which can easily lead to problems such as video stuttering, inference delay fluctuations, and analysis failures, affecting system stability. (YU & CHENG, 2022)

5.2 Technical Optimization and Improvement Strategies

In response to the core difficulties in the application of the above-mentioned technologies, and in light of the development trends of cutting-edge technologies such as artificial intelligence, big data, and edge computing, this paper proposes targeted optimization and improvement strategies. First, optimize the complex scene adaptive processing algorithm by introducing pre-processing algorithms such as image enhancement, light and shadow correction, defogging and noise reduction, and dynamic contrast adjustment. At the same time, expand the complex scene dataset and conduct scene adaptive model training to effectively improve perception accuracy in harsh environments and complex scenes and reduce recognition errors caused by environmental interference. Second, build a standardized data interaction system, unify the data transmission format, communication protocol and access standard of all-domain monitoring equipment, build a dedicated data middle platform to achieve unified access, cleaning, fusion and scheduling of multi-source heterogeneous data, completely break the data silos of equipment, and enhance the ability of all-domain situation fusion analysis. Third, introduce visual large model semantic understanding technology, combine time series big data to complete the iterative training of latent anomaly models, deeply mine the latent features of progressive, low-probability abnormal situations, strengthen the system's risk prediction ability, and achieve all-round, in-depth situation awareness and judgment in scenarios. Fourth, improve the data security and privacy protection system, adopt technical means such as video image desensitization, facial information encryption, and fuzzy processing of sensitive data, establish a data classification and grading control mechanism, standardize the full-process management of data collection, transmission, storage, and application, and eliminate the risks of privacy leakage and data abuse. 5 Build an edge-cloud collaborative computing power scheduling architecture, dynamically allocate computing power resources based on the concurrent volume of real-time video streams and the volume of analysis tasks, balance the accuracy of model inference and the real-time performance of system operation, and effectively solve the problem of system instability caused by insufficient computing power of edge devices.

6. Technical Application Scenarios and Development Trends

6.1 Core Application Scenarios

Ai-based global situational awareness technology for video surveillance has the characteristics of strong versatility, high scalability and wide adaptability scenarios, and can be widely applied in multi-domain intelligent control scenarios. In the field of urban public security, the technology can achieve real-time perception and intelligent early warning of people gathering, abnormal loitering, illegal disturbances and public security risks in key urban areas, contribute to the construction of smart security cities and improve the refinement level of urban public security control. In the field of road traffic management, it can monitor the dynamic situation of vehicle and pedestrian flow in the entire urban area, accurately identify abnormal conditions such as illegal parking, driving in the wrong direction, occupying the road, traffic congestion and traffic accidents, and provide data support for traffic diversion, order control and accident handling. In the field of security in parks, campuses and communities, it can achieve all-round situation monitoring at entrances and exits, boundary barriers and public activity areas, effectively preventing risks such as illegal intrusion of outsiders, dense gatherings of people and safety accidents, and ensuring regional security and stability. In the field of emergency control, it can monitor in real time emergencies such as fires, waterlogging on roads, people trapped, and regional dangers, quickly assess the coverage and evolution trend of the disaster, and provide scientific basis for emergency rescue command, resource allocation, and on-site handling. In the field of industrial production and construction site supervision, it can accurately identify problems such as workers' violation of operating procedures, abnormal operation of equipment, accumulation of safety hazards, and personnel absence from their posts, and achieve intelligent supervision throughout the entire process and all areas of safe production.

6.2 Future Trends

With the continuous iteration and upgrading of technologies such as visual large models, digital twins, 5G/6G communications, multimodal perception, and edge computing power scheduling, video surveillance global situation awareness technology will continue to evolve in the direction of high intelligence, high accuracy, full dimension, and high security. First, large models deeply empower situation analysis, relying on the powerful semantic understanding, logical reasoning, and autonomous learning capabilities of visual large models to break through the feature recognition limitations of traditional algorithms and achieve deep semantic analysis and intelligent decision-making in complex scenarios, significantly enhancing the intelligence level of the system. Second, the integration of digital twin technology is implemented to build a digital twin virtual model of the surveillance scene, achieving real-time mapping and synchronous update of the physical scene and the virtual situation, making the overall situation display more three-dimensional, visual and precise. Third, the deep integration of multimodal perception breaks the technical limitations of single video perception, integrates video, audio, radar, sensor, meteorological and other multimodal data to build an all-round, three-dimensional perception system, further enhancing the comprehensiveness and accuracy of situation analysis. Fourth, the system's autonomous collaboration capabilities will be upgraded. In the future, the situation awareness system will have the capabilities of autonomous learning, autonomous iteration, autonomous scheduling, and autonomous linkage, and can automatically optimize algorithm models according to scene changes, gradually realizing the operation mode of unattended and autonomous control. Fifth, standardization of security compliance system. With the continuous improvement of regulations related to personal information protection and data security, video data desensitization, privacy protection, and compliance application will become industry technical standards, achieving coordinated development of intelligent perception efficiency and data security protection.

7. Conclusion

The iterative innovation of artificial intelligence technology has completely reshaped the development pattern of the video surveillance industry, transforming video surveillance from a traditional image acquisition and data storage tool to a core intelligent governance device that integrates all-domain situation awareness, intelligent judgment, risk early warning, and proactive prevention and control, providing key technical support for the refined governance of smart cities and safe cities. This paper focuses on the core theme of AI-enabled all-domain situation awareness in video surveillance, systematically sorts out the core connotations and technical advantages of all-domain situation awareness, builds a five-layer modular all-domain situation awareness system architecture, and deeply explores core key technologies such as lightweight target detection, trajectory tracking, anomaly identification, and multi-source data fusion, The superiority of the technical solution in this paper was verified through multi-scenario comparative experiments. The core difficulties of the current technology implementation were analyzed in light of the current industry situation, and targeted optimization and improvement strategies as well as future development trends were proposed.

References

Chen, J. F., & Wu, J. Q. (2023). *Urban Security Science and Technology*, *2023*(06), 45-50.

Huang, W. J., & Zheng, K. (2023). Video dynamic situation awareness technology based on spatio-temporal deep learning. *Pattern Recognition and Artificial Intelligence*, *36*(09), 821-828.

Li, Z. X., & Zhao, K. (2024). Lightweight Video object detection and anomaly recognition algorithm based on YOLOv8. *Modern Electronics Technology*, *47*(05), 112-116.

Lin, Z. Y. (2024). Research on Privacy Security Protection Technology for Artificial Intelligence Video surveillance. *Information Security and Communication Confidentiality*, *2024*(02), 67-73.

Liu, M., & Chen, Y. (2024). *Application Research of Computers*, *41*(03), 892-896. Urban Video surveillance anomaly situation awareness based on deep learning.

Qaraqe, M., Elzein, A., & Basaran, E. (2024). PublicVision: A Secure Smart Surveillance System for Crowd Behavior Recognition. *IEEE Access*, *12*, 26474-26491.

Tian, Y., & Li, S. Y. (2024). Technical difficulties and Optimization paths of global situation awareness in intelligent video surveillance. *Intelligent Computers and Applications*, *14*(01), 189-193.

WANG, H. F., ZHANG, F., ZHANG, K. et al. (2022). Design of high performance video processingframework based on content. *Command Infornation System and Technology*, *13*(2), 85-90.

Wang, H. Y., & Zhou, M. Y. (2023). Architecture Design of Video global Situation Awareness System

Empowered by Edge Computing. *Information Technology*, *47*(08), 78-83.

YU, S. C., & CHENG, C. S. (2022). Video postback technology for military micro UAV based onsuper resolution algorithm. *Command Information System and Technology*, *13*(6), 23-28.

Zhang, L., Wang, J., & Li, H. R. (2023). Research on the application of artificial intelligence in global situation awareness of intelligent Video surveillance. *Computer Engineering and Applications*, *59*(12), 231-238.

Zhao, M., & Sun, B. (2023). Research on a Model for Global Situation Assessment of Video Surveillance based on Multi-source Data Fusion. *Big Data Technology and Application*, *11*(04), 210-216.