Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 17, No. 3, 2026

Research on Neural Network Technology in Public Safety Anomaly and Hazard Detection

Yinchan Zhu

Anhui University of Foreign Languages, Hefei, Anhui, China

Abstract: Abnormal hazards in public places have the characteristics of low frequency, diversity and strong timeliness. Single video model is easily affected by occlusion, illumination and network interruption. In this article, an abnormal danger detection method is constructed, which combines the temporal and spatial characteristics of video, environmental sensing information and communication status, and introduces the neural network synchronization control mechanism under DoS attack. The model uses convolutional network to extract local appearance and motion information, describes the evolution of events with gated time series units, and then completes multi-source feature fusion through attention mechanism. When DoS attacks cause intermittent loss of node data, state observation and synchronization compensation are introduced to maintain the time consistency between the edge detection results and the central model. The experimental results show that the F1 value of the proposed method reaches 0.922, which remains at 0.861 under the packet loss rate of 40%. Synchronous compensation reduces the average state error by about 41.7%. The research shows that the combination of dangerous content identification and network operation state can improve the stability of public safety anomaly detection under the condition of disturbed communication.

Keywords: Public Safety; Abnormal Hazard Detection; Neural Network; DoS Attack; Synchronous Control; Multi-source Fusion.

1. Introduction

Public security governance has gradually turned to real-time perception. Dangerous events account for a relatively low proportion in monitoring data, but they put forward higher requirements for identification and alarm timeliness. Traditional methods rely on manual setting of thresholds or rules, and are prone to false positives in complex scenes such as occlusion and visual angle changes. Deep neural network can learn the appearance, behavior and scene characteristics of events and become an important technical path for video anomaly detection [1-4]. The existing research mainly adopts supervised identification, normal pattern reconstruction and time series modeling. Cyclic network, three-dimensional convolution and Transformer further enhance the ability of time series feature extraction [5-7]. Related research focuses on video content, but pays little attention to data anomalies caused by multi-node collaboration and network attacks.

In the edge computing architecture, cameras and sensor nodes need to transmit features to the regional center through the network. DoS attack may lead to packet loss or communication interruption, so that old frames and missing data are misjudged as anomalies by the model. The research of networked control shows that event triggering, state observation and piecewise Lyapunov analysis can improve the synchronization performance of the system under DoS conditions [8-10], which can provide reference for the stable cooperation of public safety detection nodes.

Based on this, a neural network detection method integrating video, environmental information and communication status is proposed, and an edge node synchronization compensation module is set up. During DoS attack, the system compensates the output of the lost node by observing the state, and reduces the state drift by synchronization error constraint.

2. Public Safety Anomaly Detection Framework

2.1. Scene Characteristics and Task Definition

The judgment of public safety anomalies depends on both the degree of behavior deviation and the consequences of the scene. Therefore, the input is divided into video stream, environmental sensing stream and communication state stream, which carry visual information, environmental parameters and network indicators respectively. After the edge nodes are aligned in time, the three types of data are input into the model with 16 frames sliding window. The detection task outputs hazard probability, anomaly category and risk level for each window. Among them, anomalies include behavioral events and environmental/equipment anomalies, and the determination of comprehensive probability, duration and regional weight of risk level. For events with insufficient training, the model calculates abnormal scores according to the deviation degree of normal mode to reduce missed detection. The feature composition is shown in Table 1.

Table 1. Composition of Multi-source Input Features

Feature Stream	Representative Variables	Sampling	Primary Function
Video	RGB frames, optical flow, object tracks	25 fps	Appearance and motion encoding
Environment	Sound level, temperature, smoke, crowd density	2-10 Hz	Cross-modal evidence
Network	Packet loss, RTT, queue length, time offset	1 Hz	Reliability estimation
Context	Zone type, time period, risk weight	Event-based	Risk calibration

2.2. System Framework

The overall framework is shown in Figure 1. The front-end sensing device collects video and sensing data. Edge nodes complete denoising, target tracking and time window segmentation. The feature coding module extracts spatial, motion and short time series features respectively. The risk fusion module outputs the abnormal score in combination with the scene weight. The alarm module sends the event category, location, credibility and evidence fragments to the disposal end. The network state estimation module identifies network anomalies caused by DoS, and the synchronization controller adjusts the observation and fusion weights accordingly. After communication is resumed, nodes only upload key features and state differences to avoid the backlog of data flooding into the center in a short time.

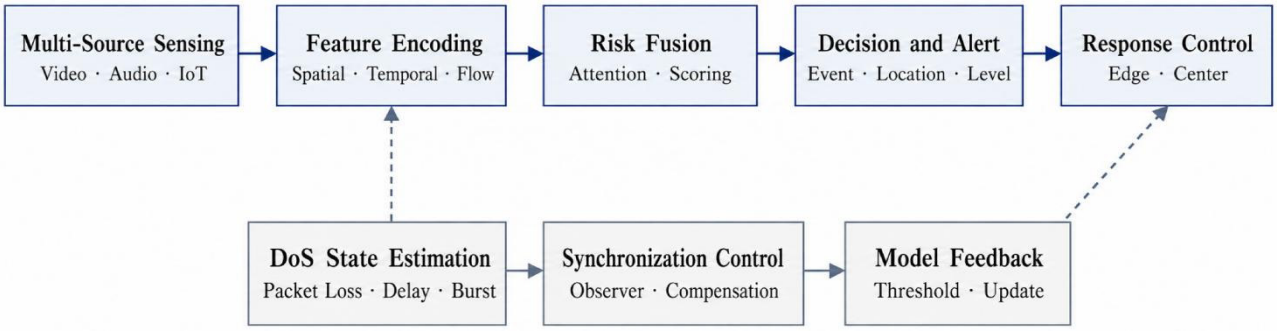


Figure 1. Detection Framework of Abnormal Danger in Public Safety

3. Multi-source Neural Network Detection and Synchronous Control Method

3.1. Spatio-temporal Feature Coding and Adaptive Fusion

The video branch extracts the multi-scale features of each frame with a lightweight convolution backbone, and adds a motion difference channel in the middle layer. Target-level features are composed of regional features in the detection frame and trajectory speed. Scene-level features describe crowd density and global motion direction. Subsequently, the bidirectional gating unit encodes the continuous window, and uses the time sequence information before and after to distinguish the instantaneous occlusion from the persistent anomaly. In the environmental branch, a two-layer fully connected network is used to process the sensor quantities with different dimensions, and the sensor offline state is indicated by the missing mask.

For the t time window, let the video, environment and network characteristics be v_t, s_t and n_t respectively. After linear mapping, the fusion is formed by nonlinear function to represent z_t; Attention coefficient α_k is dynamically allocated according to the effectiveness of each mode. The calculation form is shown in equation (1).

$$z_t = \phi\left(W_v v_t + W_s s_t + W_n n_t + b\right)$$
$$\alpha_k = \frac{\exp\left(q^{\cdot} h_k\right)}{\sum_j \exp\left(q^{\cdot} h_j\right)} \quad (1)$$

Where W_v, W_s and W_n are the learnable mapping matrices of each branch, b is the offset term, and ϕ is the ReLU activation function; h_k represents the hidden feature of k mode, q is the attention query vector, and Table α_k shows the contribution of this mode to the current judgment. When the network packet loss rate increases, n_t will reduce the weight of video features corresponding to the lost nodes, but environmental sensing and evidence of adjacent nodes can still participate in the judgment.

3.2. Joint Anomaly Score and Loss Function

The fused features enter the classification header and the normal mode reconstruction header respectively. The classification head gives the probability of known danger categories, and the reconstruction head learns the potential distribution of normal windows. When an event belongs to the existing category in the training set, the classification probability is the main weight. When the classification confidence is not high but the reconstruction error continues to increase, the system marks it as an unknown anomaly and enters the manual review queue.

The training target consists of three parts: class imbalance loss, normal sample reconstruction loss and node synchronization error, as shown in equation (2).

$$\mathrm{L} = \lambda_1 \mathrm{L}_{\mathrm{focal}}(y, p) + \lambda_2 \| x - \hat{x} \|_2^2 + \lambda_3 \sum_i e_i^{\cdot} Pe_i \quad (2)$$

In the equation, $\mathrm{L}_{\mathrm{focal}}$ alleviates the problem of rare abnormal samples by adjusting the weights of difficult samples; x and $\hat{x}$ are normal features and their reconstructed results, respectively; e_i is the synchronization error between the i th edge node and the central reference state, and p is the positive definite weight matrix; λ_1, λ_2, and λ_3 control the relative effects of three types of objectives. In the experiment, the three values were set to 1.0, 0.35, and 0.20.

3.3. State Synchronization and Compensation under DoS Attacks

DoS attacks are usually characterized by continuous packet loss or a sharp drop in available bandwidth. In this article, $\delta_i(t)$ represents the communication state of node i at time t: $\delta_i(t) = 0$ when the link is available, and $\delta_i(t) = 1$ when it is attacked or blocked. In normal state, nodes are updated according to the state difference of adjacent nodes; When communication is interrupted, the observer gives the

predicted value according to the pre-attack state, local sensing change and historical evolution trend. The control input is written as Equation (3).

$$u_i(t) = [1 - \delta_i(t)]K \sum_j a_{ij}\left[x_j(t) - x_i(t)\right] + \delta_i(t)K_h\left[\hat{x}_i(t) - x_i(t)\right] \tag{3}$$

Where $u_i(t)$ is the node state correction, a_{ij} is the connection weight in the network topology, K is the synchronization gain in the normal communication phase, K_h is the compensation gain in the holding phase, x_i and x_j are the node i and its neighboring node states respectively, and $\hat{x}_i(t)$ is the observer prediction state. The first one uses neighborhood differences to promote consistency when the link is available, and the second one suppresses the unconstrained drift of node state during DoS.

Figure 2 shows the typical response under three-segment DoS interference. Without synchronous control, the state error quickly accumulates after the attack starts, and it still takes a long time to fall back after the communication is restored. After adding observation compensation, the peak error and recovery time are obviously reduced. The function of synchronous control is to ensure that the danger detection link still has continuous output ability in short-term communication blocking.

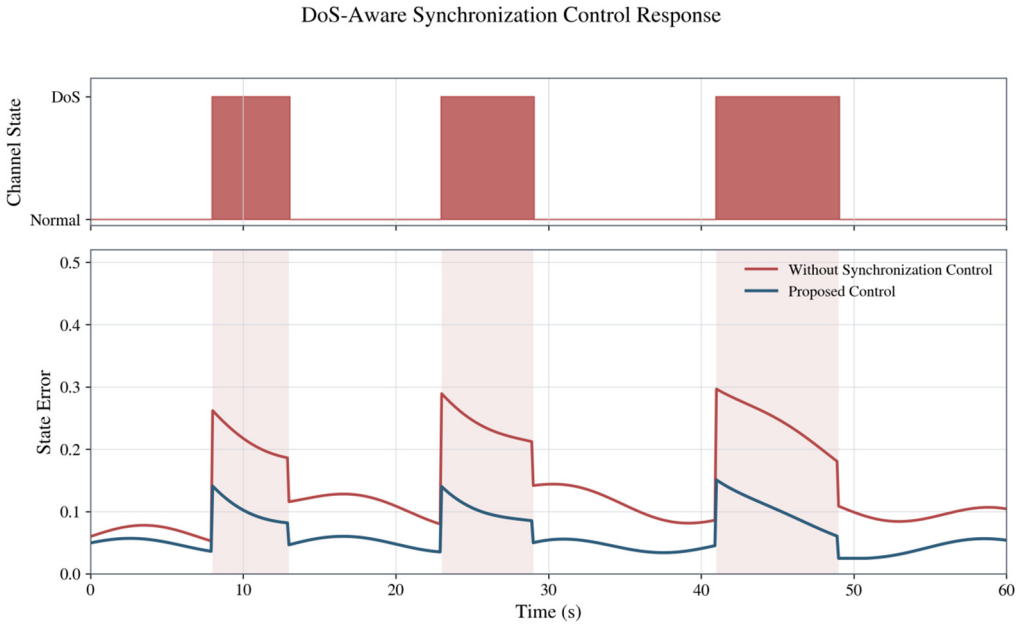


Figure 2. Synchronization Control Response under DoS Attack

4. Experimental Design and Result Analysis

4.1. Data and Experimental Setup

The experimental video uses publicly available video clips from UCSD Ped2, CUHK Avenue, and ShanghaiTech anomaly detection data, and is uniformly input at 25 fps and 640 × 360 resolution according to public space monitoring tasks. The training set focuses on normal behavior, while the test set includes events such as running, going against the flow, conflicts, and abnormal gatherings. To supplement the scenes of fireworks and environmental changes, temperature, smoke, and sound pressure sequences are added while retaining the original video annotations. On the communication side, based on the edge nodes uploading features once per second, set a random packet loss of 0% to 40% and a burst block of 3 to 8 seconds.

The model adopts the AdamW optimizer, with an initial learning rate of 1 × 10-4, a batch size of 16, and 60 rounds of training. The video window contains 16 frames, updated every 0.32 seconds. The experiment selected 3D-CNN, CNN-GRU, and asynchronous compensation models for comparison, mainly evaluating detection performance, inference delay, and synchronization state error. The threshold remains unchanged after being determined by the validation set. The experimental environment and parameters are shown in Table 2.

Table 2. Experimental Environment and Main Parameters

Item	Configuration	Item	Configuration
GPU	NVIDIA RTX 4070	Framework	PyTorch 2.3
Input size	640 × 360	Window length	16 frames
Batch size	16	Training epochs	60
Optimizer	AdamW	Initial LR	1×10^{-4}
Packet loss	0%-40%	DoS burst	3-8 s

4.2. Detection Performance and Comparison Results

The comprehensive test results are shown in Table 3. 3D-CNN has limited cross-frame modeling ability, and its F1 value is 0.838. After CNN-GRU introduced time series

coding, the recognition of continuous anomalies was improved. The F1 value of this method is 0.922 and AUC is 0.954, which is 4.2 percentage points higher than that of CNN-GRU. After the synchronization compensation is removed, the F1 value of the model in the mixed DoS scene drops to 0.870. This shows that the data mismatch caused by communication disturbance will affect the fusion judgment.

Table 3. Anomaly Detection Performance of Different Methods

Method	Accuracy	Precision	Recall	F1-score	AUC	Latency (ms)
3D-CNN	0.862	0.851	0.826	0.838	0.884	34.7
CNN-GRU	0.891	0.887	0.873	0.880	0.921	41.9
Proposed w/o Sync	0.883	0.879	0.861	0.870	0.913	35.2
Proposed	0.926	0.918	0.927	0.922	0.954	38.6

The model has a good recognition effect on violent conflict and rapid escape. The error of abnormal aggregation mainly comes from normal queuing and short-term congestion. After adding environmental sensing features, the false alarm rate of pyrotechnics in low light conditions decreased from 8.6% to 4.9%. Residue detection is more susceptible to occlusion and short-term stay.

4.3. Robustness under DoS Disturbance

Figure 3 shows the comprehensive performance of different methods and the robustness under the condition of packet loss. When the packet loss rate increases to 40%, the F1 value of this method decreases from 0.922 to 0.861, while that of the unsynchronized control model decreases to 0.731. The gap between them increases with the increase of packet loss rate.

In the burst blocking scenario, synchronization compensation reduces the average state error by about 41.7% and shortens the convergence time after communication recovery to 2.6 s. Under the condition of 40% packet loss, the end-to-end alarm delay is 0.83 s, which is still lower than the preset response target of 1 s.

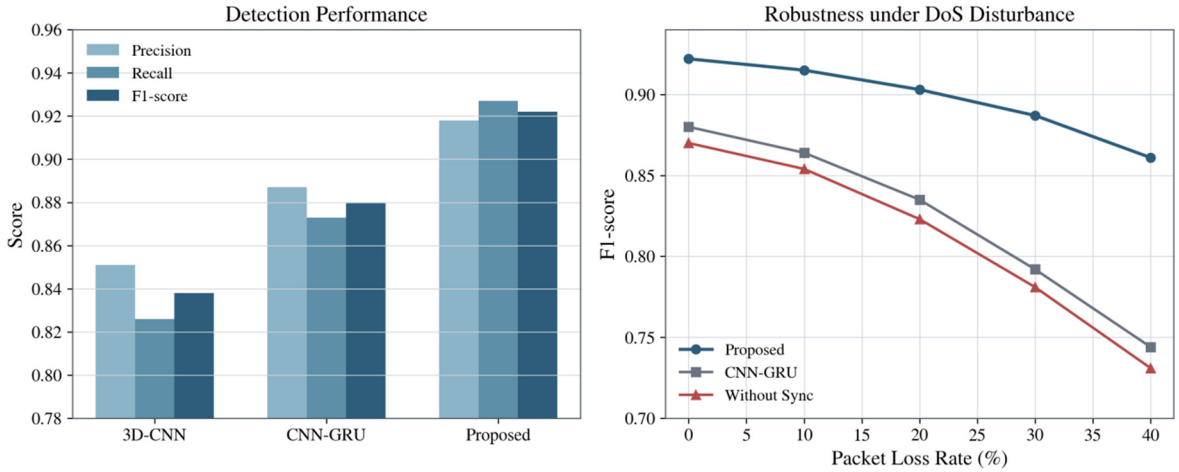


Figure 3. Detection Performance and DoS Disturbance Robustness

5. Discussion

The stability of public safety anomaly detection is not only affected by the accuracy of visual model, but also depends on whether multi-node evidence can be aligned in time. When communication is disturbed, the mixing of old and new features may cause false time series changes. It is helpful to reduce such false positives by taking the network state as reliability information and maintaining the node state stable during DoS.

The function of multi-source fusion is related to the type of events. Environmental sensing information is more helpful for events such as fireworks. Its gain on violence or legacy is relatively limited. Communication status only reflects the integrity of evidence and cannot be directly used as a danger criterion. The system should output the content risk score and communication reliability respectively, and keep the source of evidence.

The effect of synchronous compensation is also affected by attack duration and local observation quality. Long-term communication interruption will reduce the prediction reliability, so the automatic disposal authority can be gradually reduced according to the interruption duration. The network side still needs to combine link redundancy and traffic protection measures. Synchronous compensation is mainly used to ensure the continuity of detection.

This article verifies the logic of the method based on open data and simulated communication disturbance, but there is still a gap with the real city-level application. Follow-up can carry out long-term evaluation of cross-regional and multi-equipment, and introduce uncertainty estimation and online calibration.

6. Conclusion

In this article, a multi-source neural network detection framework is constructed around the problems in the detection of abnormal dangers in public safety. Methods Environmental sensing information was used to correct hazard semantics, and the reliability of evidence was evaluated by communication status. Aiming at the node loss caused by DoS attack, the observation state and synchronization error constraints are introduced. The results show that the F1 value and AUC of the proposed method are 0.922 and 0.954, and the F1 value of 0.861 is still maintained at the packet loss rate of 40%. Synchronous compensation reduces the average state error by 41.7% and controls the alarm delay within 0.83 s. It can be seen that the combination of anomaly recognition model and synchronous control under attack conditions can improve the reliability of public safety awareness system in complex network environment. In practical application, it is still necessary to calibrate the

threshold according to the regional risk, equipment capacity and disposal process.

Acknowledgments

This work is supported by the University-level Key Scientific Research Project of Anhui University of Foreign Languages: Research on Neural Network Synchronous Control Algorithm under DoS Attacks (Project No. Awky 2025047).

References

[1] Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. ACM Computing Surveys, 54(2), 1-38. https://doi.org/10.1145/3439950.

[2] Nayak, R., Pati, U. C., & Das, S. K. (2021). A comprehensive review on deep learning-based methods for video anomaly detection. Image and Vision Computing, 106, 104078.

[3] Gandapur, M. Q. (2022). E2E-VSDL: End-to-end video surveillance-based deep learning model to detect and prevent criminal activities. Image and Vision Computing, 123, 104467. https://doi.org/10.1016/j.imavis.2022.104467.

[4] Ingle, P. Y., & Kim, Y. G. (2022). Real-time abnormal object detection for video surveillance in smart cities. Sensors, 22(10), 3862. https://doi.org/10.3390/s22103862.

[5] Bajgoti, A., Gupta, R., Balaji, P., & Dwivedi, R. (2023). SwinAnomaly: Real-time video anomaly detection using Video Swin Transformer and SORT. IEEE Access, 11, 111094-111105. https://doi.org/10.1109/ ACCESS. 2023. 332 1801.

[6] Huszár, V. D., Adhikarla, V. K., Négyesi, I., & Krasznay, C. (2023). Toward fast and accurate violence detection for automated video surveillance applications. IEEE Access, 11, 18772-18793. https://doi.org/10.1109/ACCESS.2023.3245521.

[7] Kotkar, V. A., & Sucharita, V. (2023). Fast anomaly detection in video surveillance system using robust spatiotemporal and deep learning methods. Multimedia Tools and Applications, 82, 34259-34286. https://doi.org/10.1007/s11042-023-14840-0.

[8] Huang, X., Ding, D. W., Li, Z., & An, C. (2022). Distributed security control for complex cyber-physical systems against denial-of-service attacks. International Journal of Control, Automation and Systems, 20(2), 421-431. https://doi.org/ 10. 1007/ s12555-020-0906-7.

[9] Liu, J., Shen, M., & Xie, X. (2023). Secure consensus control for multi-agent systems subject to consecutive asynchronous denial-of-service attacks. International Journal of Control, Automation and Systems, 21, 2214-2226. https://doi.org/ 10. 1007/s12555-021-0564-4.

[10] Wen, G., Wang, P., Lv, Y., Chen, G., & Zhou, J. (2023). Secure consensus of multi-agent systems under denial-of-service attacks. Asian Journal of Control, 25(2), 695-709. https://doi. org/ 10. 1002/asjc.2953.