A Framework for Intelligent Campus Security: Integrating Pedestrian Re-identification into Management Systems

Feiyang Wang

School of Cyberspace Security, Chengdu University of Information Technology (Xingu Industrial College), Chengdu 610225, China

Abstract: ***Existing campus surveillance systems mostly rely on facial recognition, which fails in scenarios involving long distances, occlusions, and backlighting. To address this issue, this paper introduces person re-identification technology, leveraging body-pose features for identification, and designs and implements a campus security management system that overcomes the shortcomings of facial recognition. The system captures video via multiple RTSP cameras, accurately detects pedestrians with YOLOv8, and feeds the detected images into a person re-identification model to extract high-dimensional features such as clothing, posture, and gait for cross-camera matching and continuous tracking. Deployed on Alibaba Cloud GPU-accelerated servers, it adopts a cloud-client separation architecture: the cloud provides alerts and low-latency video streams, while the client delivers visualized monitoring, real-time localization, and intelligent alarms. The data layer uses MySQL to store basic information and event logs, Redis to cache feature vectors for faster retrieval, and reserves a high-spec GPU scaling plan. The system achieves high-precision, multi-camera continuous tracking under complex surveillance angles, compensating for the limitations of traditional facial recognition and enhancing campus-wide security.***

Keywords: Campus security system; Person re-identification; Pose feature recognition; YOLOv8; Motion trajectory analysis.

1. INTRODUCTION

Campus security, as a core component of the smart-city public-safety system, directly affects the lives and property of students and staff as well as social stability. Current mainstream campus-security systems rely mainly on facial-recognition technology, using access control, visitor registration, and behavior monitoring to improve management efficiency. However, in complex surveillance scenarios such as long distances, occlusion, backlighting, or low illumination, the failure rate of capturing facial features rises sharply, causing recognition to fail. Meanwhile, large-scale collection and storage of facial data faces strict constraints under the Data Security Law and the Personal Information Protection Law, with risks of data leakage and misuse. Field investigations reveal that existing systems in dynamic, complex environments suffer from unclear imaging, untimely behavior prediction, limited functionality, and slow manual response. To address these issues, this paper introduces person re-identification technology, extracting full-body pose features of pedestrians and fusing multi-source visual signals such as clothing information, body shape, and gait to achieve identity association and trajectory tracking across camera views. Its core advantages are weak dependence on facial data, cross-domain tracking capability, and privacy-friendly design. Recommendation systems research is significantly advanced by Wang (2025) through joint training of propensity and prediction models using targeted learning for handling data missing not at random [1], while interactive systems progress through Xie and Chen's (2025) InVis framework for human-centered data interpretation [2]. Digital advertising technologies show substantial innovation with Zhu's (2025) RAID system for reliability automation in large-scale ad platforms [3], Zhang's (2025) CrossPlatformStack enabling high-availability deployment across meta services [4], Hu's (2025) GenPlayAds for procedural playable 3D ad creation [5], and Li, Lin, and Zhang's (2025) privacy-preserving framework incorporating federated learning and differential privacy [6]. Urban computing and public infrastructure benefit from Xu's (2025) CivicMorph for generative public space modeling [7], while communication systems advance through Tu's (2025) SmartFITLab for intelligent 5G interoperability testing [8]. Data analytics is enhanced by Xie and Liu's (2025) DataFuse for multimodal interview analytics [9], and workflow automation transforms through Zhu's (2025) TaskComm for small business efficiency [10] and Zhang's (2025) reinforcement learning approach to ad campaign optimization [11]. Content creation is revolutionized by Hu's (2025) few-shot neural editors for 3D animation [12], while industrial applications encompass Tan's (2024) analysis of AI trends in automotive production [13]. Digital transformation extends to Zhuang's (2025) theoretical construction of real estate marketing strategies [14], and recommendation systems further evolve through Han and Dou's (2025) hierarchical graph attention networks with multimodal knowledge graphs [15]. Healthcare applications include Yang's (2025) Prompt-Biomrc model for intelligent consultation [16] and conversational AI advances through Yang et al.'s (2025) RLHF fine-tuning for alignment with implicit user feedback [17], complemented by parallel optimization methods for LLM-based recommendation systems [18]. Business intelligence features Zhang et al.'s (2025) AI-driven sales forecasting in gaming [19], while web technologies advance through Yang's (2025) website optimization using Dijkstra's algorithm [20]. Urban planning benefits from Xu's (2025) UrbanMod for accelerated city architecture planning [21], healthcare through Hsu et al.'s (2025) MEDPLAN for personalized medical plans [22], and cross-media analytics through Yuan and Xue's (2025) fusion framework [23]. Computer vision includes Chen et al.'s (2022) gaze estimation research [24], medical diagnostics through Wang's (2025) RAGNet for arthritis risk prediction [25], and natural language processing through Yu et al.'s (2025) automatic text summarization research [26].

2. DESIGN OF A CAMPUS SECURITY MANAGEMENT SYSTEM BASED ON PERSON RE-IDENTIFICATION TECHNOLOGY

2.1 Terminal Data Collection

The system selects the TP-Link Tapo C210 network camera as the data-collection terminal [4]. Protocol support: native compatibility with RTSP (Real Time Streaming Protocol) [5]. The camera continuously pushes RTSP video streams to the server via the campus LAN.

2.2 Server-Side Person Re-identification

The server side is deployed on an Alibaba Cloud Linux cloud server equipped with GPUs, offering strong parallel-computing power suitable for high-performance model training and real-time inference. To stably ingest and distribute multi-camera video streams, the server runs the rtsp-simple-server component, which is based on the RTSP standard and handles RTSP stream ingestion and distribution. Together with FFmpeg, it performs video decoding, format conversion, and stream redirection, ensuring that video from cameras of different encodings and resolutions can be stably and efficiently ingested in the cloud for downstream algorithms.

2.2.1 Pedestrian Detection Training

During the pedestrian detection phase, the system employs Ultralytics YOLOv8 (You Only Look Once, a single-shot detection framework [8]) for model training. The training data structure is composed of both public datasets and self-collected datasets, which exhibit rich diversity across multi-camera, complex occlusion, and lighting variation scenarios. The self-collected data covers typical campus scenes such as teaching-building corridors, cafeteria entrances, and sports-field exits to enhance the model’s adaptability to real deployment environments. Spatial transformations and color perturbations are applied during training to augment data diversity, and a cosine annealing strategy dynamically adjusts the learning rate, combined with batch normalization and cross-GPU synchronized training to optimize convergence efficiency. Throughout training, mAP (mean Average Precision) serves as the core monitoring metric to ensure model detection accuracy and robustness.

2.2.2 Person Re-identification

In the person re-identification phase, the system uses the Torchreid library to build the feature extraction network [3]. During data preparation, public data and campus-collected samples are organized into separate folders per identity, forming a ReID dataset containing hundreds of thousands of images. The network backbone is OSNet (Omni-Scale Network), which remains lightweight while incorporating channel attention mechanisms to extract more discriminative feature vectors. Training employs a joint optimization strategy combining cross-entropy loss and triplet loss, along with BNneck (Batch Normalization Neck) to regularize feature vectors and reduce distribution discrepancies caused by different camera viewpoints. Hyperparameters include an initial learning rate of 0.00035, a batch size of 64 (16 per GPU), the Adam optimizer, and freezing the backbone layers in later training stages to fine-tune only the fully connected layers, thereby improving the model’s transferability to new scenes.

2.2.3 Model Deployment and Real-time Inference

During deployment, the server continuously receives multiple RTSP video streams and uses OpenCV to parse frames. After YOLOv8 performs inference on each frame, pedestrian images are cropped according to detection boxes and fed into the re-identification model to extract fixed-length feature vectors. Extracted vectors are first compared rapidly in the Redis (Remote Dictionary Server) vector cache; if similarity exceeds the set threshold, the corresponding identity is returned directly. Otherwise, the full identity features stored in the MySQL database are invoked for precise matching, and new identity vectors are added to the Redis cache on demand to optimize subsequent queries. The system implements asynchronous inference and multi-threaded scheduling, enabling parallel execution of detection and recognition pipelines and significantly reducing per-frame latency.

2.3 Frontend Design

This system uses the Python Qt framework binding PyQt (also called Python Qt) for GUI development. The graphical user interface supports remote control and visualization of detection results. The interface includes a video display window, start/stop detection buttons, a camera-switch button, and a recognition result list. After development, the application is packaged into a Windows executable with the one-click tool PyInstaller, making deployment simple and efficient.

2.3.1 Multi-protocol communication between client and server

Control commands: transmitted via HTTP (HyperText Transfer Protocol) and HTTPS (HyperText Transfer Protocol Secure) interfaces to adjust service status in real time. Detection results: pushed to the front end as JSON (JavaScript Object Notation) data through WebSocket, dynamically updating bounding boxes and person IDs. Video stream: delivered via WebRTC (Web Real-Time Communication) and rendered with low latency through Qt WebEngine or the aiortc library.

2.3.2 Front-end recognition display and modular design

After receiving the JSON data pushed by WebSocket, the front end parses the bounding-box coordinates and person IDs, then uses the Qt drawing interface or the OpenCV open-source computer vision library to overlay recognition boxes on the video and label target identities, while displaying person IDs and suspicious status in an information list.

2.3.3 The client provides three core functions

First, unknown persons identified by the system are marked as suspicious and trigger an alert; operators can upload an image of the target, and the system returns the person's real-time location. Second, for abnormal behaviors such as prolonged loitering or gathering in the same area, the system sends an alert to the front end. Third, for detected strangers, the front-end interface offers a trusted-person registration feature.

2.4 System Database

In the campus security management system based on person re-identification, the database layer deploys both Redis and MySQL services on an Alibaba Cloud GPU-accelerated Linux server. The Redis cluster handles real-time caching of person re-identification feature vectors and vector-index retrieval. Intelligent expiration and priority eviction rules are set for each vector to balance memory utilization and retrieval performance. The MySQL database stores structured business data, including a campus personnel registration table, an abnormal behavior event table, and a multi-camera trajectory association table. The personnel registration table records fields such as student ID, name, college, authorized areas, and feature-vector index. The event table aggregates timestamps, camera IDs, and associated feature vectors for suspicious gathering alarms, illegal intrusion alerts, and loitering warnings. The trajectory association table links the same identity across different viewpoints chronologically, enabling visual analysis and trace playback of pedestrian flow on campus. Through primary-foreign key relationships and partitioned table design, the system supports second-level behavior statistics, regional heat-map generation, and focused-group tracking queries.

2.5 System Security Mechanism

This project, aimed at campus security management, has built a multi-layered security protection system.

1) Interface authentication uses the JSON Web Token mechanism; all REST APIs and WebSocket connections related to pedestrian recognition, alert management, and user registration must carry a valid signed token and pass server-side asymmetric-encryption verification before access is granted.

2) End-to-end data transmission is protected by TLS 1.2 encryption [9], securing image streams and metadata from campus monitoring nodes to the cloud server.

3) The MySQL database is configured with a least-privilege policy, allowing only the monitoring-center application account to perform insert, delete, update, and select operations.

4) Redis is deployed in a dedicated Virtual Private Cloud subnet and isolated via access-control lists and an internal whitelist.

2.6 System Deployment

This campus safety management project runs pedestrian detection and re-identification services, the database, and streaming services together on an Alibaba Cloud GPU-accelerated Linux instance. The deployment process includes: 1) Basic environment setup—install Python and its dependencies; 2) Model and configuration distribution—upload models to object storage and enable automatic sync and update on startup; 3) Database initialization—run SQL scripts to create tables, build Redis vector indexes, and preload features; 4) Streaming service configuration—deploy an RTSP streaming service to connect cameras and configure reverse proxy and SSL (Secure Sockets Layer) certificates; 5) Detection and recognition service startup—launch YOLOv8 and ReID inference processes in parallel, leveraging GPU for real-time processing; 6) Backend API and alert push—start the WebSocket service and RESTful API and configure secure communication channels; 7) Client release and access—package the PyQt client to enable video preview, annotation, alerts, registration, and trajectory functions, while supporting automatic updates and backend connectivity.

3. SUMMARY

Centered on pedestrian re-identification technology, the system integrates key algorithms such as YOLOv8 object detection, deep-feature matching, cross-camera identity matching, and trajectory tracking to build an intelligent security management system tailored for campus safety. By deploying RTSP-compatible video capture terminals and combining PyTorch-trained YOLOv8 and ReID deep neural network models, it achieves precise detection and multi-angle identity recognition of human targets. The system overcomes the failure of traditional facial recognition in complex scenes, enabling high-accuracy pedestrian recognition and trajectory tracking based on posture and appearance features, thereby improving campus security response efficiency and proactive defense capabilities.

FUNDING:

This work is supported by the Undergraduate Education Teaching Research and Reform Project and Undergraduate Teaching Engineering Project of Chengdu University of Information Technology, No. JYJG2022135, No. JYJG2024169, and the Undergraduate Innovation Training Program “Campus Security Management System Based on Person Re-identification (No. X202510621276).”

REFERENCES

[1] Wang, Hao. "Joint Training of Propensity Model and Prediction Model via Targeted Learning for Recommendation on Data Missing Not at Random." AAAI 2025 Workshop on Artificial Intelligence with Causal Techniques. 2025.
[2] Xie, Minhui, and Shujian Chen. "InVis: Interactive Neural Visualization System for Human-Centered Data Interpretation." Authorea Preprints (2025).
[3] Zhu, Bingxin. "RAID: Reliability Automation through Intelligent Detection in Large-Scale Ad Systems." (2025).
[4] Zhang, Yuhan. "CrossPlatformStack: Enabling High Availability and Safe Deployment for Products Across Meta Services." (2025).
[5] Hu, Xiao. "GenPlayAds: Procedural Playable 3D Ad Creation via Generative Model." (2025).
[6] Li, X., Lin, Y., & Zhang, Y. (2025). A Privacy-Preserving Framework for Advertising Personalization Incorporating Federated Learning and Differential Privacy. arXiv preprint arXiv:2507.12098.
[7] Xu, Haoran. "CivicMorph: Generative Modeling for Public Space Form Development." (2025).

[8] Tu, Tongwei. "SmartFITLab: Intelligent Execution and Validation Platform for 5G Field Interoperability Testing." (2025).

[9] Xie, Minhui, and Boyan Liu. "DataFuse: Optimizing Interview Analytics Through Multimodal Data Integration and Real-Time Visualization." (2025).

[10] Zhu, Bingxin. "TaskComm: Task-Oriented Language Agent for Efficient Small Businesses Workflows." (2025).

[11] Zhang, Yuhan. "Learning to Advertise: Reinforcement Learning for Automated Ad Campaign Optimization for Small Businesses." (2025).

[12] Hu, Xiao. "Learning to Animate: Few-Shot Neural Editors for 3D SMEs." (2025).

[13] Tan, C. (2024). The Application and Development Trends of Artificial Intelligence Technology in Automotive Production. Artificial Intelligence Technology Research, 2(5).

[14] Zhuang, R. (2025). Evolutionary Logic and Theoretical Construction of Real Estate Marketing Strategies under Digital Transformation. Economics and Management Innovation, 2(2), 117-124.

[15] Han, X., & Dou, X. (2025). User recommendation method integrating hierarchical graph attention network with multimodal knowledge graph. Frontiers in Neurorobotics, 19, 1587973.

[16] Yang, J. (2025, July). Identification Based on Prompt-Biomrc Model and Its Application in Intelligent Consultation. In Innovative Computing 2025, Volume 1: International Conference on Innovative Computing (Vol. 1440, p. 149). Springer Nature.

[17] Yang, Zhongheng, Aijia Sun, Yushang Zhao, Yinuo Yang, Dannier Li, and Chengrui Zhou. "RLHF Fine-Tuning of LLMs for Alignment with Implicit User Feedback in Conversational Recommenders." arXiv preprint arXiv:2508.05289 (2025).

[18] Yang, Haowei, Yu Tian, Zhongheng Yang, Zhao Wang, Chengrui Zhou, and Dannier Li. "Research on Model Parallelism and Data Parallelism Optimization Methods in Large Language Model-Based Recommendation Systems." arXiv preprint arXiv:2506.17551 (2025).

[19] Zhang, Jingbo, et al. "AI-Driven Sales Forecasting in the Gaming Industry: Machine Learning-Based Advertising Market Trend Analysis and Key Feature Mining." (2025).

[20] Yang, Yifan. "Website Internal Link Optimization Strategy and SEO Effect Evaluation Based on Dijkstra Algorithm." Journal of Computer, Signal, and System Research 2.3 (2025): 90-96.

[21] Xu, Haoran. "UrbanMod: Text-to-3D Modeling for Accelerated City Architecture Planning." Authorea Preprints (2025).

[22] Hsu, Hsin-Ling, et al. "MEDPLAN: A Two-Stage RAG-Based System for Personalized Medical Plan Generation." arXiv preprint arXiv:2503.17900 (2025).

[23] Yuan, Yuping, and Haozhong Xue. "Cross-Media Data Fusion and Intelligent Analytics Framework for Comprehensive Information Extraction and Value Mining." (2025).

[24] Chen, J., Zhang, X., Wu, Y., Ghosh, S., Natarajan, P., Chang, S. F., & Allebach, J. (2022). One-stage object referring with gaze estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5021-5030).

[25] Wang, Y. (2025). RAGNet: Transformer-GNN-Enhanced Cox–Logistic Hybrid Model for Rheumatoid Arthritis Risk Prediction.

[26] Yu, Z., Sun, N., Wu, S., & Wang, Y. (2025, March). Research on Automatic Text Summarization Using Transformer and Pointer-Generator Networks. In 2025 4th International Symposium on Computer Applications and Information Technology (ISCAIT) (pp. 1601-1604). IEEE.