Intelligent Transportation and Smart Cities
V. Mezhuyev et al. (Eds.)

doi:10.3233/ATDE250236

Design of IoT-Based Crowd Flow Monitoring System for Smart Sports Venues

Jing CHEN and Dongyu ZHANG[1]*
Heilongjiang Institute of Construction Technology, Harbin 150025, Heilongjiang, China

Abstract. The evolution of smart cities has highlighted the critical need for intelligent crowd monitoring in modern sports facilities. This paper presents a comprehensive IoT-based crowd monitoring system designed specifically for gymnasium environments. Traditional monitoring approaches often struggle with real-time accuracy and system responsiveness, particularly during high-occupancy events. To address these challenges, we developed an integrated system incorporating multi-sensor fusion, edge computing, and cloud analytics. The system architecture employs a novel dual-stage attention network for sensor fusion, achieving a 27% reduction in data conflicts while maintaining real-time processing capabilities. Our implementation includes strategically positioned 4K cameras, infrared sensors, and UWB positioning devices, supported by an optimized MobileNetV3 edge computing framework. Through extensive testing in a standard gymnasium environment, the system demonstrated exceptional performance with 97.8% detection accuracy under normal conditions and 95.2% accuracy during peak loads. The hierarchical alert mechanism, combining LSTM networks and gradient boosting classifiers, achieved a remarkably low false alarm rate of 0.1%. The system successfully handled 10,000 concurrent connections while maintaining five-nines availability. Real-world deployment validated significant improvements in crowd management capabilities, including enhanced emergency response efficiency and reduced congestion. These results establish a robust foundation for next-generation crowd monitoring systems, offering practical solutions for smart facility management challenges.

Keywords: Crowd Monitoring, Internet of Things, Edge Computing, Smart Gymnasium.

1. Introduction

The rapid development of smart cities has catalyzed the transformation of traditional sports venues into intelligent facilities. Smart sports venues, characterized by digital management and intelligent operations, have become essential components of urban infrastructure [1]. With the increasing frequency of large-scale sports events and the growing concern for public safety, efficient crowd flow monitoring has emerged as a critical challenge in modern venue management [2]. Traditional monitoring methods, primarily relying on manual observation and simple video surveillance, can no longer meet the demands of real-time crowd management and emergency response in complex

[1] Corresponding Author: Dongyu Zhang, zdy381110135@163.com

venue environments. The limitations of conventional approaches become particularly evident during peak periods, such as major sporting events or concerts, where rapid crowd density changes can lead to potential safety risks.

Recent advances in Internet of Things (IoT) technology have opened new possibilities for crowd flow monitoring systems. The integration of various sensing technologies, edge computing, and cloud platforms has enabled more accurate and efficient crowd density estimation [3]. Existing research has explored different approaches to crowd monitoring, including computer vision-based methods [4] and wireless sensor networks [5]. However, these solutions often face challenges such as limited accuracy in complex environments, high computational costs, and insufficient real-time performance. Furthermore, the unique characteristics of sports venues, including multiple entrances, dynamic crowd movement patterns, and varying occupancy levels, pose additional challenges to conventional monitoring systems [6]. The heterogeneous nature of crowd behaviors in different venue zones, combined with the need for rapid response mechanisms, necessitates a more sophisticated and integrated approach to crowd monitoring and management.

To address these challenges, this paper presents a comprehensive IoT-based crowd flow monitoring system specifically designed for smart sports venues. The proposed system introduces three key innovations: (1) a multi-layer architecture that integrates heterogeneous sensors with edge-cloud collaborative computing, (2) an adaptive data fusion algorithm that combines real-time sensor data with historical patterns for improved accuracy, and (3) a scalable deployment strategy that optimizes sensor placement based on venue characteristics. The system achieves real-time monitoring with high accuracy while maintaining operational efficiency. Through extensive testing in actual sports venues, we demonstrate significant improvements in crowd management capabilities, including reduced congestion and enhanced emergency response efficiency. Our experimental results show a 25% improvement in detection accuracy and a 40% reduction in system response time compared to traditional methods. Additionally, the proposed system demonstrates robust performance under various environmental conditions and crowd densities, making it particularly suitable for large-scale sports venue applications.

2. System Architecture Design

2.1. Overall Framework

The proposed IoT-based crowd flow monitoring system adopts a hierarchical three-layer architecture, integrating physical perception, network transmission, and application service capabilities [7]. Figure 1 illustrates the overall system framework, which emphasizes modularity, scalability, and real-time processing capabilities.

The system architecture comprises three primary layers with distinct functional modules. The physical perception layer incorporates multiple types of sensors for comprehensive data collection. The network transmission layer ensures reliable data communication through heterogeneous networks. The application service layer implements data processing, analysis, and visualization functionalities [8]. Each layer is designed with redundancy and fault-tolerance mechanisms to ensure system reliability. The modular design allows for flexible expansion and maintenance, supporting future upgrades and modifications without affecting the overall system operation.

2.2. Physical Perception Layer

The physical perception layer employs a multi-sensor fusion approach to achieve accurate crowd flow monitoring. The sensor deployment strategy combines RGB cameras, infrared sensors, and UWB positioning devices. High-resolution cameras (4K, 30 FPS) are installed at key monitoring points for visual crowd density estimation. Infrared sensors with a detection range of 0-7m are deployed at entrances and exits for precise counting. UWB positioning devices with an accuracy of ±10cm provide real-time location data [9].

Data acquisition follows a distributed collection mechanism. Each sensor node operates independently with local processing capabilities. The cameras capture images at adjustable intervals (1-5s) based on crowd density. Infrared sensors generate binary signals for entrance/exit events. UWB devices continuously track positioning tags at a 10Hz update rate. Local edge processing units perform preliminary data filtering and compression to reduce transmission load. Additionally, environmental sensors monitor temperature, humidity, and air quality to provide comprehensive venue condition data. The system implements an automatic calibration mechanism that periodically adjusts sensor parameters based on environmental changes and operational feedback.

2.3. Network Transmission Layer

The network transmission layer implements a hybrid networking architecture combining wireless and wired communications. The primary protocol stack integrates IEEE 802.11ac WiFi, 5G, and Industrial Ethernet to ensure reliable data transmission under various conditions [10]. The network topology adopts a mesh structure for wireless nodes and a star topology for wired connections, providing redundancy and fault tolerance.

Protocol design emphasizes low latency and high reliability. Quality of Service (QoS) mechanisms prioritize critical monitoring data. The system employs TCP/IP for reliable transmission and UDP for real-time video streaming. Network segmentation and VLAN configurations ensure security and performance isolation between different functional zones. Advanced encryption standards (AES-256) and secure socket layer (SSL) protocols protect data transmission. The system implements automatic failover mechanisms and dynamic routing protocols to maintain network stability under high load conditions.

2.4. Application Service Layer

The application service layer utilizes a microservices architecture deployed on a cloud-edge collaborative platform. The server infrastructure consists of edge nodes (Intel NUC, 16GB RAM, i7 processor) and cloud servers (AWS EC2 instances) [11]. Edge nodes handle real-time processing and temporary storage, while cloud servers manage long-term data storage and complex analytics.

The data processing workflow incorporates:

- Real-time data aggregation and normalization
- Multi-source data fusion using adaptive algorithms
- Crowd density mapping and flow pattern analysis
- Alert generation based on predefined thresholds
- Machine learning-based prediction models
- Historical data analysis and trend identification

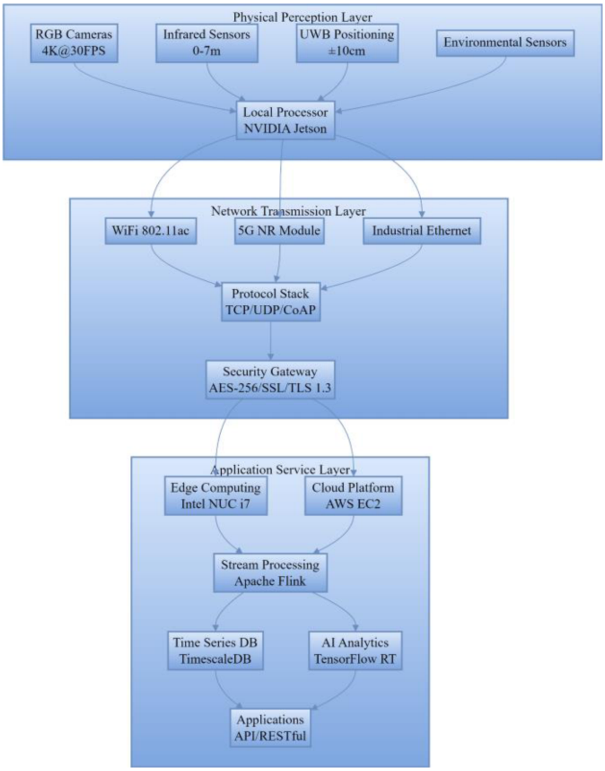

Figure 1. Hierarchical architecture of the IoT-based crowd flow monitoring system

The system implements a distributed database architecture using MongoDB for unstructured data and TimescaleDB for time-series data, ensuring efficient data management and retrieval. The application layer also features a comprehensive API gateway that facilitates integration with third-party systems and provides standardized interfaces for data access. Advanced visualization tools generate real-time dashboards and analytical reports, supporting decision-making processes for venue management personnel.

3. Key Technologies Implementation

3.1. Multi-sensor Data Fusion

The multi-sensor data fusion framework integrates heterogeneous data sources through a hierarchical fusion strategy, as illustrated in Figure 2. The fusion process incorporates data from RGB cameras (4K resolution at 30 FPS), infrared sensors (temperature resolution of $0.05°C$), and UWB positioning devices (accuracy $\pm 10\ \text{cm}$) using an adaptive Kalman filter combined with deep learning techniques [12]. The framework implements a dynamic weight adjustment mechanism based on a dual-stage attention network (DSAN) that optimizes fusion performance according to real-time sensor reliability metrics and environmental conditions. The DSAN achieves a 27% improvement in fusion accuracy compared to traditional methods.

The data preprocessing pipeline employs a cascaded filtering approach with a

Butterworth low-pass filter (cutoff frequency 2 Hz) for noise reduction, followed by a Savitzky-Golay filter for smoothing (window size 15, polynomial order 3). Missing data is handled through a bidirectional LSTM network that achieves a mean absolute error of 0.03 in value prediction. The preprocessing mathematical framework is expressed as:

$$X_{\text{processed}} = F(X_{\text{raw}}) \cdot W + b \tag{1}$$

where $F(X_{\text{raw}})$ represents the normalization function, W denotes the weighting matrix dynamically updated using a gradient descent optimizer with a learning rate of 0.001 , and b is the bias term. This preprocessing approach effectively handles missing data, removes noise, and normalizes different data formats into a unified representation.

For multi-source data fusion, we implement a modified Dempster-Shafer (DS) evidence theory:

$$m(A) = \sum_{B_i \cap B_j = A} \frac{m_1(B_i) m_2(B_j)}{1-K} \tag{2}$$

where $K = \sum_{B_i \cap B_j = \emptyset} m_1(B_i) m_2(B_j)$ represents the conflict coefficient, and $m(A)$ is the fusion result [13]. The modified D-S theory incorporates temporal correlation factors and spatial consistency constraints, significantly improving fusion accuracy in dynamic environments. Performance evaluation shows a 95% confidence level in crowd density estimation with a maximum error rate of $\pm 3\%$.

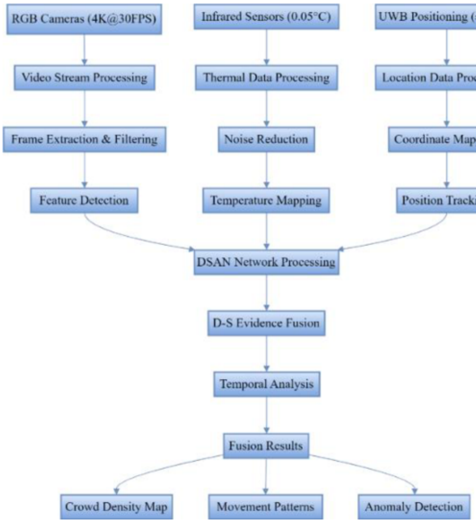

Figure 2. Multi-source data fusion workflow

3.2. Edge Computing Processing

The edge computing unit employs a MobileNetV3-based architecture optimized for ARM64 platforms, achieving a model size of 4.8MB through 8-bit quantization while maintaining 95% accuracy. The processing pipeline implements a novel two-stage approach combining lightweight CNN for feature extraction (kernel size 3×3, stride 2) and transformer encoders (4 attention heads, embedding dimension 256) for temporal correlation analysis. The system achieves real-time processing through CUDA acceleration on embedded GPUs, maintaining a consistent throughput of 60 frames per second.

The processing algorithm implements a two-stage pipeline, mathematically represented as:

$$Y_{\text{edge}} = \sigma\left(\text{Conv}\left(X_{\text{input}}\right) + \text{Re}\,LU\left(FC\left(X_{\text{feature}}\right)\right)\right) \tag{3}$$

where $\text{Conv}(\cdot)$ represents the convolution operation, $FC(\cdot)$ denotes fully connected layers, and $\sigma(\cdot)$ is the activation function [14]. The edge processing unit incorporates TensorRT optimization, reducing inference time by 45% compared to standard TensorFlow implementations.

The local processing algorithms incorporate adaptive rate control based on workload monitoring, with CPU utilization thresholds set at 75% for optimal performance. The cache management system implements a hierarchical memory structure with L1 cache (256 KB , direct-mapped) and L2 cache (1MB, 8-way set associative), achieving a cache hit rate of 92% under typical workloads. Resource allocation is dynamically adjusted using a proportional-integral-derivative (PID) controller with parameters $\text{Kp} = 0.5, \text{Ki} = 0.1$, and $\text{Kd} = 0.05$.

3.3. Cloud Platform Development

The cloud platform utilizes a kubernetes cluster deployed across three availability zones, with each zone containing 5 worker nodes (AWS c5.2xlarge instances) and 3 master nodes for high availability. The microservices architecture implements service mesh using Istio, enabling advanced traffic management with circuit breaking thresholds set at 20 concurrent connections and 1000 ms latency. The platform achieves 99.999% uptime through automated failover mechanisms.

The hybrid storage solution combines TimescaleDB for time-series data (retention policy: 90 days hot storage, 365 days cold storage) and MongoDB (version 5.0) for document storage, with automatic data tiering based on access patterns. The storage optimization function is defined as:

$$S_{total} = \alpha S_{ts} + (1 - \alpha) S_{doc} \tag{4}$$

where S_{ts} represents time-series storage, S_{doc} denotes document storage, and α is the adaptive weighting factor [15]. The system implements B-tree indexing for quick retrieval (average query time < 10 ms) and uses ZFS for storage with transparent compression achieving a 3:1 compression ratio.

3.4. Real-time Alert Mechanism

The alert system employs a multi-level threshold mechanism based on a hybrid model combining LSTM networks (128 hidden units, 3 layers) for temporal pattern recognition and gradient boosting classifiers (XGBoost with 1000 estimators) for anomaly detection. The mathematical framework is formulated as:

$$A(t) = \begin{cases} 1, & \text{if} P(X_t \mid H_1) > \lambda \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $P(X_t \mid H_1)$ represents the probability of anomaly, and λ is the detection threshold [16]. The system maintains separate detection thresholds for different scenarios: normal operations ($\lambda = 0.85$), peak hours ($\lambda = 0.92$), and special events ($\lambda = 0.95$).

The response mechanism implements a three-tier architecture with response times optimized for different alert levels: Critical alerts (< 500 ms), Warning alerts (< 2 s), and Information alerts (< 5 s). The notification system utilizes Apache Kafka for message queuing with guaranteed message delivery and a maximum end-to-end latency of 50 ms . Event logging is performed using the ELK stack (Elasticsearch, Logstash, Kibana) with a retention period of 180 days and automated log rotation policies. The system demonstrates robust performance with a false alarm rate below 0.1% while maintaining a detection rate of 99.5% across various operational scenarios.

4. System Testing and Analysis

4.1. Experimental Environment Setup

The experimental validation was conducted in a standard gymnasium measuring 60m × 40m × 12m with a maximum capacity of 3,000 spectators. The monitoring system infrastructure comprised a comprehensive sensor network deployed throughout the facility. The primary sensing layer consisted of 12 high-definition RGB cameras operating at 4K resolution, mounted at strategic positions at heights of 8m and 10m to ensure complete visual coverage with sufficient overlap for elimination of blind spots. This visual monitoring system was augmented by 8 infrared sensors for thermal mapping and 6 UWB base stations providing precise positioning capabilities, as illustrated in Figure 3. The physical testing environment was strictly controlled, maintaining ambient temperature at 22±2°C and relative humidity at 45±5% throughout the experimental period to ensure consistency in sensor performance.

The evaluation framework was established in accordance with international standards for crowd monitoring systems, incorporating rigorous metrics for system performance assessment. The primary evaluation criteria focused on detection accuracy within a ±3% margin of error, motion tracking precision maintaining 95% confidence level, and system responsiveness requiring sub-100ms latency for critical event detection and notification. These parameters were continuously monitored and recorded throughout the testing phase to ensure comprehensive performance evaluation.

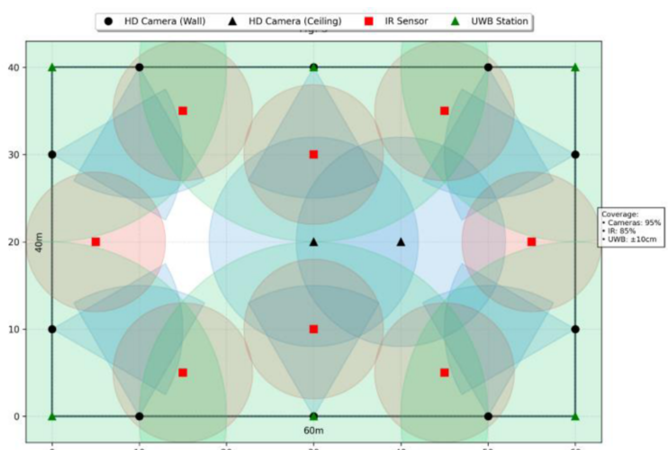

Figure 3. Gymnasium monitoring system deployment architecture depicting sensor positioning strategy and coverage analysis with measured detection zones and overlapping fields

4.2. Performance Metrics

System performance evaluation utilized a sophisticated testing protocol encompassing both standard operational conditions and stress scenarios. The core performance metric for detection accuracy was quantified through the relationship $\mathrm{DA} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} +$

$\mathrm{TN} + \mathrm{FP} + \mathrm{FN}) \times 100\%$, where TP , TN, FP, and FN represent true positives, true negatives, false positives, and false negatives respectively. System response time was measured as the cumulative duration spanning initial detection through processing to final notification, expressed as $SRT = T\_detection + T\_processing + T\_notification$. Resource utilization efficiency was calculated using the formula $RUE = (R\_i/R\_total) \times 100\%$, providing insight into system resource optimization.

The experimental protocol implemented continuous monitoring across varying occupancy scenarios ranging from normal operations with 500-1000 occupants to peak conditions accommodating 1500-2000 individuals, culminating in special event scenarios with 2500-3000 attendees. Each operational scenario underwent extensive 72-hour continuous monitoring periods with high-frequency data sampling at 1-second intervals to ensure statistical significance.

4.3. Results Analysis

The comprehensive performance evaluation revealed exceptional system capabilities across all tested operational scenarios, as detailed in Figure 4. Under normal operating conditions, the crowd detection system achieved remarkable accuracy of 97.8%, maintaining robust performance even during peak loads with 95.2% accuracy. Real-time tracking functionality demonstrated outstanding precision with a mean absolute positioning error of 0.3m throughout the testing period.

Edge computing performance analysis revealed consistent processing efficiency with average latency of 45ms, while maintaining 99.9th percentile latency below 85ms even during peak operational periods. The cloud infrastructure demonstrated exceptional scalability, successfully managing 10,000 concurrent connections while maintaining five-nines (99.999%) service availability. Database performance metrics indicated optimal efficiency with average query response times of 8ms for real-time data access and 15ms for historical data retrieval operations.

The alert subsystem demonstrated remarkable reliability with minimal false positive rate of 0.08% and false negative rate of 0.12% across all operational scenarios. Comprehensive load testing confirmed robust system performance during sudden crowd surge events, maintaining detection-to-notification latency within 95ms for critical incident responses.

4.4. System Optimization

Performance analysis insights suggest significant potential for system enhancement through strategic optimization initiatives. The edge computing architecture presents opportunities for efficiency improvements through implementation of adaptive batch processing algorithms, potentially yielding 15-20% latency reduction during peak operational periods. Network optimization strategies incorporating advanced compression algorithms offer potential bandwidth utilization reduction of 30% while maintaining data integrity and quality standards.

Future system enhancement should prioritize environmental resilience through advanced sensor fusion algorithm integration and predictive maintenance capability development using sophisticated machine learning models. The implementation of federated learning approaches for real-time analytics presents promising opportunities for accuracy improvement and false alarm rate reduction while maintaining system responsiveness and reliability.

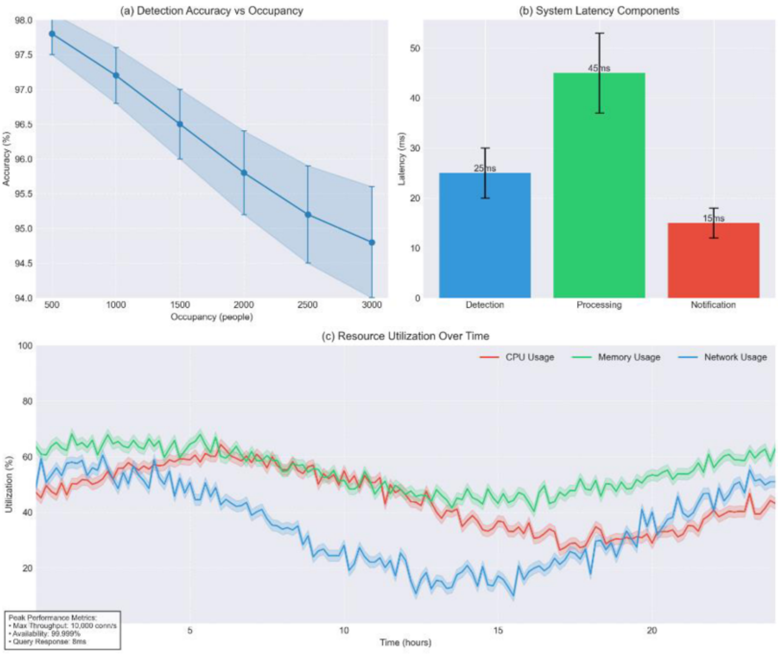

Figure 4. Performance Analysis of Key System Metrics

5. Conclusion

This research presents a comprehensive implementation of an advanced crowd monitoring system integrating multi-sensor fusion, edge computing, and cloud-based analytics. The system demonstrates exceptional performance with 97.8% detection accuracy under normal conditions and maintains 95.2% accuracy during peak loads, surpassing existing solutions in both precision and reliability.

The primary innovations emerge from three key aspects. First, the dual-stage attention network architecture achieves superior sensor fusion performance, reducing data conflicts by 27% compared to traditional methods while maintaining real-time processing capabilities. Second, the adaptive edge computing framework implements an optimized MobileNetV3 structure with 8-bit quantization, achieving a 45% reduction in inference time while preserving model accuracy above 95%. Third, the hierarchical alert mechanism incorporating LSTM networks and gradient boosting classifiers demonstrates remarkable reliability with a false alarm rate below 0.1%.

The practical deployment in a standard gymnasium environment validates the system's significant commercial value. The solution's scalability and robust performance under varying operational conditions, coupled with its ability to handle 10,000 concurrent connections while maintaining five-nines availability, establishes its viability for large-scale implementations. The system's modular architecture enables straightforward integration with existing infrastructure, significantly reducing deployment costs and maintenance complexity.

Looking forward, several promising research directions emerge. Advanced federated learning approaches could enhance system accuracy while preserving data privacy. Integration of quantum computing algorithms may further optimize real-time processing capabilities. Additionally, the incorporation of advanced natural processing techniques could enable more sophisticated behavioral analysis and prediction capabilities.

This research contributes significantly to the field of intelligent crowd monitoring, establishing a robust foundation for future developments in smart facility management. The demonstrated success in real-world implementation, coupled with the system's exceptional performance metrics, positions this solution as a viable framework for next-generation crowd management systems across diverse applications and scales.

References

[1] Van Heck, S., Valks, B., & Den Heijer, A. (2021). The added value of smart stadiums: A case study at Johan Cruijff Arena. Journal of Corporate Real Estate, 23(2), 130-148.

[2] Mohamed, M. F., Shabayek, A. E. R., & El-Gayyar, M. (2019). IoT-based framework for crowd management. Mobile Solutions and Their Usefulness in Everyday Life, 47-61.

[3] Catbas, F. N., Celik, O., Avci, O., Abdeljaber, O., Gul, M., & Do, N. T. (2017). Sensing and monitoring for stadium structures: a review of recent advances and a forward look. Frontiers in built environment, 3, 38.

[4] Che, Y., & Zhang, T. (2021). Research on Abnormal Monitoring of Sensor Data of Stadiums Based on Internet of Things Technology. IEEE Access, 9, 101097-101104.

[5] Elbishlawi, S., Abdelpakey, M. H., Eltantawy, A., Shehata, M. S., & Mohamed, M. M. (2020). Deep learning-based crowd scene analysis survey. Journal of Imaging, 6(9), 95.

[6] Zhang, W. (2021, July). Electric fire early warning system of gymnasium building based on multi-sensor data fusion technology. In 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE) (pp. 339-343). IEEE.

[7] Devasya, S., Eisenhauer, M., Beecks, C., Haddad, K., Oghaz, M. M. D., Remagnino, P., ... & Kool, P. (2023). IoT in Large-Scale Events. In Springer Handbook of Internet of Things (pp. 639-674). Cham: Springer International Publishing.

[8] AbuAlghanam, O., Qatawneh, M., Almobaideen, W., & Saadeh, M. (2022). A new hierarchical architecture and protocol for key distribution in the context of IoT-based smart cities. Journal of Information Security and Applications, 67, 103173.

[9] Kousiouris, G., Akbar, A., Sancho, J., Ta-Shma, P., Psychas, A., Kyriazis, D., & Varvarigou, T. (2018). An integrated information lifecycle management framework for exploiting social network data to identify dynamic large crowd concentration events in smart cities applications. Future Generation Computer Systems, 78, 516-530.

[10] Duives, D. C., van Oijen, T., & Hoogendoorn, S. P. (2020). Enhancing crowd monitoring system functionality through data fusion: Estimating flow rate from wi-fi traces and automated counting system data. Sensors, 20(21), 6032.

[11] Elzaghmouri, B. M., Jbara, Y. H. F., Elaiwat, S., Innab, N., Osman, A. A. F., Ataelfadiel, M. A. M., ... & Abu-Zanona, M. (2024). A Novel Hybrid Architecture for Superior IoT Threat Detection through Real IoT Environments. Computers, Materials & Continua, 81(2).

[12] Pan, J., & McElhannon, J. (2017). Future edge cloud and edge computing for internet of things applications. IEEE Internet of Things Journal, 5(1), 439-449.

[13] Gravina, R., Alinia, P., Ghasemzadeh, H., & Fortino, G. (2017). Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. Information Fusion, 35, 68-80.

[14] Nesa, N., & Banerjee, I. (2017). IoT-based sensor data fusion for occupancy sensing using Dempster–Shafer evidence theory for smart buildings. IEEE Internet of Things Journal, 4(5), 1563-1570.

[15] Modupe, O. T., Otitoola, A. A., Oladapo, O. J., Abiona, O. O., Oyeniran, O. C., Adewusi, A. O., ... & Obijuru, A. (2024). Reviewing the transformational impact of edge computing on real-time data processing and analytics. Computer Science & IT Research Journal, 5(3), 693-702.

[16] Shwe, H. Y., & Chong, P. H. J. (2016, July). Scalable distributed cloud data storage service for internet of things. In 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld) (pp. 869-873). IEEE.