An IoT-Enabled Intelligent Data Center Monitoring Framework Using YOLO-Based Image Analysis for Green Cities

Yousef Farhaoui [1,*], Ahmad El Allaoui [1], Hamed Taherdoost [2], and Bharat Bhushan [3]

[1] IMIA Laboratory, T-IDMS Research Team, Faculty of Sciences and Techniques of Errachidia (FSTE), Moulay Ismail University of Meknès, Morocco
[2] School of Arts, Science and Technology, University Canada West, Vancouver, Canada
[3] Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Jharkhand, India
Email: y.farhaoui@fste.umi.ac.ma (Y.F.); hmad666@gmail.com (A.E.A.); hamed.taherdoost@gmail.com (H.T.); bharat_bhushan1989@yahoo.com (B.B.)
*Corresponding author

***Abstract*—This study presents an intelligent data center monitoring and surveillance framework for green cities, integrating Internet of Things (IoT) technology with You Only Look Once (YOLO)-based object detection. The proposed system enhances data center security and operational efficiency by enabling real-time identification of critical components such as servers, routers, cooling systems, and unauthorized personnel. Leveraging computer vision and deep learning, the framework provides proactive monitoring and rapid response to potential threats. The YOLO model is trained on a large, annotated dataset specifically curated for data center environments, allowing precise detection and localization of IoT-based objects of interest. Additional modules incorporating thermal imaging and motion detection further strengthen anomaly recognition and intrusion prevention capabilities. Performance evaluation across multiple data center scenarios demonstrates high detection accuracy, low latency, and strong robustness under varying environmental conditions. The study also highlights the computational efficiency achieved through GPU parallelization and model optimization, ensuring real-time processing in large-scale deployments. Overall, the proposed framework offers a scalable, energy-efficient, and secure solution for smart data center management, contributing to the advancement of sustainable and resilient green city infrastructures.**

***Keywords*—Internet of Things (IoT), You Only Look Once (YOLO) object detection, smart data centers, green cities, deep learning, real-time monitoring, edge computing**

I. Introduction

Data centers constitute the backbone of modern digital infrastructure, enabling cloud computing, big data analytics, artificial intelligence services, and global connectivity [1]. With the rapid expansion of smart cities and green urban ecosystems, ensuring secure, efficient, and sustainable data center operations has become a critical priority [2]. These infrastructures must be continuously monitored to guarantee operational reliability, energy efficiency, physical security, and regulatory compliance [3].

Recent advances in the Internet of Things (IoT) and deep learning have opened new opportunities for intelligent monitoring and surveillance in data center environments [4]. IoT devices enable the acquisition of real-time data from diverse sources, including cameras, environmental sensors, and energy meters, while deep learning models offer powerful capabilities for automated analysis and decision-making [5]. Object detection models based on the You Only Look Once (YOLO) family have demonstrated remarkable performance in real-time visual monitoring tasks due to their high accuracy and low latency [6–8].

In this context, this paper proposes an advanced IoT-enabled intelligent monitoring framework for data centers in green cities, leveraging the YOLOv9 deep learning architecture for real-time object detection and analysis [9]. The proposed system focuses on accurately detecting and classifying critical components and events, including servers, networking equipment, cooling systems, and unauthorized human presence. By enhancing situational awareness, the framework strengthens preventive security measures, supports efficient facility management, and contributes to sustainable data center operations [10–12].

Despite these advancements, several challenges remain. Data center environments are complex and dynamic, characterized by variable lighting conditions, spatial constraints, high equipment density, and strict reliability requirements [13–21]. Furthermore, large-scale deployments require scalable architectures that can handle massive data streams while maintaining low latency [22– 24]. Integrating human operators into automated monitoring systems also remains essential,

Manuscript received August 23, 2025; revised October 14, 2025; accepted December 25, 2025; published June 22, 2026.

particularly for validation, decision-making in ambiguous scenarios, and continuous system improvement [25–27].

The main objectives of this paper are therefore:

(1) To design a scalable and IoT-enabled monitoring framework tailored for data center environments in green cities [28];
(2) To evaluate the effectiveness of the YOLOv9 model for real-time object detection under diverse operational conditions [29, 30];
(3) To integrate human-in-the-loop mechanisms to enhance system reliability, adaptability, and accountability [31–33];
(4) To assess the system's performance in terms of accuracy, detection speed, robustness, and sustainability [33, 34].

II. Related Work

A. Internet of Things-Based Monitoring in Data Centers

Several studies have explored IoT-driven solutions for monitoring data centers, with a focus on environmental control, energy optimization, and fault detection [35]. IoT sensors are commonly used to track temperature, humidity, power consumption, and airflow, enabling predictive maintenance and energy-efficient management [36]. However, traditional IoT monitoring systems often lack advanced visual intelligence, limiting their ability to detect physical security threats or abnormal human activities in real time [37].

B. Deep Learning and YOLO-Based Object Detection

Deep learning-based object detection has evolved rapidly, with the YOLO family emerging as one of the most effective real-time detection frameworks [38]. From YOLOv1 to recent versions such as YOLOv8 and YOLOv9, continuous improvements have focused on detection accuracy, inference speed, and robustness under challenging conditions [39]. YOLOv9 introduces enhanced feature extraction mechanisms and optimized gradient learning strategies, making it particularly suitable for complex indoor environments such as data centers. Recent research has applied YOLO-based models to intelligent surveillance, smart buildings, and industrial monitoring. However, relatively few studies have specifically addressed data center monitoring using state-of-the-art YOLO architectures within the context of green cities and sustainable computing [40].

C. Data Center Security and Intelligent Surveillance

Security in data centers is a critical concern due to the sensitive nature of hosted data and digital services [41]. Existing approaches typically rely on access control systems, conventional CCTV surveillance, and rule-based alert mechanisms [42]. While these systems are effective to some extent, they often require extensive human supervision and lack intelligent automation for proactive threat detection [43]. Vision-based surveillance systems powered by deep learning provide a promising alternative by enabling real-time detection of unauthorized access, abnormal behavior, and infrastructure misuse [44].

D. Privacy Protection and Human-in-the-Loop Systems

Privacy and data protection are fundamental requirements in modern data center monitoring systems, especially in smart city environments governed by strict regulatory frameworks [45]. Recent works emphasize privacy-by-design principles, secure data storage, encrypted communication, and anonymization mechanisms [46]. In parallel, human-in-the-loop approaches have gained increasing attention as a means of balancing automation with accountability and ethical oversight [47].

Human operators contribute to dataset annotation, system validation, and continuous learning processes, improving robustness and reducing false detections [48]. By combining automated object detection with human supervision and feedback, intelligent IoT-based monitoring systems can achieve higher reliability, adaptability, and compliance with ethical and regulatory standards [49, 50].

III. Methodology

A. Problem Statement and System Overall Architecture

Modern data centers operate in highly dynamic and complex environments, managing massive volumes of heterogeneous data generated by servers, networking devices, cooling systems, and human activities. Ensuring real-time monitoring, security, and operational efficiency in such environments poses significant challenges, particularly when relying on conventional rule-based or manual surveillance systems. These approaches often lack scalability, adaptability, and the ability to respond proactively to abnormal events. The core problem addressed in this work is the design of an intelligent, scalable, and real-time monitoring framework capable of accurately detecting critical objects and events within data center environments while maintaining robustness under variable operational conditions. This includes the detection of infrastructure components (e.g., servers, switches, cooling units), monitoring human presence and activities, and supporting decision-making through automated analysis. To address these challenges, we propose an IoT-enabled monitoring architecture integrated with a YOLOv9-based deep learning object detection module. The overall system architecture consists of four main layers:

(1) The data acquisition layer is composed of IoT sensors and camera systems deployed across different zones of the data center;
(2) The data transmission layer is responsible for secure and low-latency communication between edge devices and processing units;
(3) The intelligent processing layer, where the YOLOv9 model performs real-time object detection and analysis;
(4) The application layer provides visualization, alerting, and decision-support functionalities for operators. This layered architecture ensures modularity, scalability, and efficient integration of human-in-the-loop mechanisms.

B. Dataset and Preprocessing

A robust and well-structured dataset is fundamental for developing reliable Artificial Intelligence (AI)-based monitoring systems in data center environments. In this study, particular attention is given to dataset composition, diversity, and environmental variability to ensure model generalization and real-world applicability [51]. The dataset is composed of annotated images representing all relevant object categories present in a data center, including servers, network switches, cooling units, power cables, and human operators. Images were collected across multiple operational zones within the facility to capture different spatial configurations and infrastructure layouts. To enhance dataset diversity, data acquisition was performed under varying lighting conditions, multiple camera viewpoints, and different equipment arrangements. Environmental variability was explicitly considered during dataset construction. This includes illumination changes due to operational fluctuations, variations in human activity levels, and temperature differences captured through complementary sensing modalities such as thermal imaging. Incorporating these factors allows the trained model to maintain stable performance under realistic and dynamic conditions. For reproducibility and transparency, detailed documentation of dataset sources, annotation guidelines, and labeling tools is provided. The annotation process followed standardized bounding box strategies and clearly defined object class labels, with quality control procedures applied to minimize labeling errors. Key training parameters are also reported, including learning rate, batch size, optimizer configuration, training epochs, and loss functions. To further strengthen model robustness, a comprehensive data preprocessing and augmentation pipeline was applied. This includes geometric transformations (rotation, scaling, and flipping), random illumination adjustments, and noise injection techniques, which collectively enhance dataset variability and reduce overfitting. A summary of the dataset characteristics, object distribution, and preprocessing strategies is presented in Table I.

TABLE I. DATASET COMPOSITION AND DIVERSITY

Object Category	Images	Annotation Format	Lighting Variations	Thermal Conditions	Occlusions	Equipment Layouts
Servers	12,000	Bounding boxes	Low/Medium/High	Yes	Yes	Rack positions
Routers	5500	Bounding boxes	Low/Medium/High	Yes	Partial	Rack positions
Cooling Units	3000	Bounding boxes	Low/Medium/High	Yes	No	Floor layouts
Personnel (Unauthorized)	4500	Bounding boxes	Low/Medium/High	Yes	Partial	Variable

C. YOLOv9 Model and Optimization for Data Centers

Performance assessment and optimization are critical components of intelligent data center monitoring systems, as they directly impact operational efficiency, resilience, and security. In the proposed framework, the YOLOv9 model serves as the core visual intelligence engine, optimized to meet the stringent real-time and reliability requirements of data center environments. YOLOv9 offers advanced feature extraction and programmable gradient learning capabilities that enhance detection accuracy while maintaining low inference latency. To adapt the model to data center-specific scenarios, targeted optimization strategies were applied, including fine-tuning on a domain-specific dataset, adjustment of detection thresholds, and anchor optimization to reflect the spatial characteristics of equipment layouts [2]. These adaptations enable accurate recognition of infrastructure components and human presence under varying illumination and spatial configurations. Security-aware optimization plays a decisive role in sustaining model performance. By preventing unauthorized access and malicious interference, security mechanisms ensure the integrity and availability of monitoring data. Network-level protections such as firewalls, Intrusion Detection Systems (IDS), and Intrusion Prevention Systems (IPS) continuously monitor traffic flows, filtering anomalous or malicious packets before they can disrupt system operations [40]. Complementary endpoint protections, including antivirus and antimalware solutions, further safeguard processing nodes against malicious code and runtime exploits. Regular patch management and software updates are systematically applied to eliminate known vulnerabilities, contributing to stable and uninterrupted model execution. Continuous monitoring and logging are integrated into the YOLOv9 optimization workflow. Infrastructure monitoring tools and Security Information and Event Management (SIEM) platforms provide real-time insights into inference latency, resource utilization, and security events. This visibility enables early detection of performance bottlenecks or anomalous behaviors, allowing proactive optimization and efficient resource allocation.

D. Multimodal Integration: Thermal Imaging and Motion Detection Modules

To enhance robustness and situational awareness, the proposed framework incorporates multimodal sensing by integrating thermal imaging and motion detection modules alongside RGB-based YOLOv9 object detection. Thermal cameras capture temperature distributions across equipment and monitored spaces, enabling early identification of overheating components or abnormal thermal patterns that may indicate hardware malfunction or safety risks. Motion detection modules complement visual object detection by continuously tracking dynamic changes within monitored zones. This allows the system to identify unexpected human movements, unauthorized access attempts, or abnormal activity patterns, even in low-light or visually occluded conditions. The fusion of RGB, thermal, and motion-based information improves detection accuracy, reduces false positives, and strengthens system reliability in complex operational environments. Multimodal data fusion is performed at the

edge level to minimize latency and bandwidth usage. By correlating thermal anomalies, motion cues, and visual detections, the system generates richer contextual insights and supports more informed decision-making for data center operators.

E. Privacy Protection and General Data Protection Regulation (GDPR) Compliance Design

To ensure lawful, ethical, and responsible deployment, a comprehensive privacy protection strategy compliant with the GDPR is embedded into the system design. The framework adheres to fundamental GDPR principles, including data minimization, purpose limitation, transparency, and secure processing of sensitive information. To reduce privacy risks, the system prioritizes on-device inference through edge computing, allowing object detection and anomaly analysis to be performed locally on IoT nodes without transmitting raw visual data to centralized servers. All communications among sensors, edge devices, and cloud components are protected using end-to-end encryption protocols, preventing unauthorized access or data manipulation. For long-term storage and analytics, detected events and metadata undergo secure anonymization procedures, ensuring that no personally identifiable information is retained beyond operational needs. The architecture follows privacy-by-design principles by integrating access control policies, encrypted logging, and restricted data retention periods directly into the system lifecycle. These measures ensure regulatory compliance while maintaining scalability for large-scale data center and smart city deployments.

F. Experimental Setup and Evaluation Metrics

The experimental evaluation was conducted to assess the effectiveness, robustness, and efficiency of the proposed YOLOv9-based monitoring framework. Experiments were performed using a combination of edge and cloud resources to reflect real-world deployment scenarios. The system was evaluated under varying lighting conditions, equipment densities, and levels of human activity to test adaptability and resilience. Performance metrics include detection accuracy, precision, recall, and mean Average Precision (mAP) to quantify object detection performance. Real-time efficiency was evaluated through inference latency and Frames-Per-Second (FPS) measurements, while resource utilization metrics such as CPU/GPU usage and memory consumption were monitored to assess scalability. In addition, robustness was analyzed by measuring performance degradation under environmental variability and simulated anomaly scenarios. Security and reliability were also considered in the evaluation process. Logging and monitoring tools were used to track system stability, anomaly detection rates, and recovery times following simulated faults or intrusion attempts. This comprehensive evaluation framework ensures that the proposed system meets both functional and non-functional requirements for intelligent data center monitoring.

IV. Results and Discussion

A. Overall Performance and Model Comparison

The primary objective of this project is to develop a real-time IoT-based monitoring and surveillance system for data centers, leveraging the advanced YOLOv9 model for high-precision object detection and tracking. The system is designed to accurately identify and localize critical components, including servers, routers, cooling units, and unauthorized individuals, within complex data center environments. By enabling proactive detection of potential threats or anomalies, the framework enhances security, operational reliability, and overall infrastructure safety. Built upon state-of-the-art computer vision and deep learning techniques, YOLOv9 is employed due to its transformer-based backbone, enhanced feature extraction capabilities, and multi-scale detection performance, which collectively provide superior accuracy and real-time responsiveness in IoT-driven scenarios [2]. The model is fine-tuned using a domain-specific dataset comprising diverse images from multiple data center environments, ensuring robustness to variations in lighting, occlusions, and environmental conditions. To further strengthen monitoring efficiency, the system integrates infrared imaging and motion detection modules. Infrared sensors enable the detection of abnormal thermal patterns, which may indicate hardware malfunctions or overheating, while motion detection modules provide immediate alerts upon identifying suspicious activities or unauthorized movements. For large-scale deployment, the system emphasizes scalability and computational efficiency. YOLOv9 is optimized for real-time inference using GPU-based parallel processing and edge computing architectures, minimizing latency while maximizing throughput. Its resilience to dynamic operational conditions ensures continuous and autonomous monitoring in mission-critical data center operations. Overall, the proposed YOLOv9-driven IoT framework provides a comprehensive, adaptive, and intelligent solution for data center surveillance, effectively combining next-generation deep learning with advanced sensing modalities to enable proactive threat detection, anomaly prevention, and operational excellence (see Table II for runtime benchmarks).

TABLE II. Runtime Benchmarks

Hardware	FPS (Inference)	Latency Per Frame (ms)	Notes
GPU (RTX 3090)	45	22	Full data center images
CPU (i9-12900)	15	66	Edge processing
Edge Device	10	100	Lightweight IoT setup

Data centers play a crucial role in storing and processing vast amounts of sensitive information, making their security and continuous monitoring essential to prevent unauthorized access and potential threats. This

study proposes a robust real-time monitoring system that leverages computer vision and deep learning for efficient detection and tracking of key objects within data centers. The system employs the YOLOv9 model, chosen for its high detection accuracy and computational efficiency in IoT-based environments. The model is trained using transfer learning with pre-trained ImageNet weights to accelerate convergence and enhance feature extraction and then fine-tuned to adapt to the specific visual conditions and operational environments of data centers. In deployment, the YOLOv9 model processes real-time video streams captured by surveillance cameras, identifying and localizing objects of interest through bounding boxes and class probabilities. This facilitates precise and rapid object detection, enabling proactive monitoring and anomaly detection. To further enhance surveillance performance, the framework integrates thermal imaging and motion detection technologies. Thermal cameras monitor temperature fluctuations to detect potential equipment malfunctions or overheating, while motion detection modules analyze video frames to identify unusual activity and trigger instant alerts. The system is designed for robustness under varying environmental conditions commonly encountered in data centers. During training, the YOLOv9 model is exposed to diverse lighting scenarios and obstacle variations to improve adaptability.

Additionally, GPU-based parallel computing ensures efficient processing and scalability for large-scale deployments. Extensive experiments were conducted to evaluate the proposed framework, measuring key performance indicators such as detection precision, processing latency, and resilience to environmental disturbances. Comparative analyses with existing monitoring approaches demonstrate the superiority of the YOLOv9-based system in terms of accuracy, speed, and robustness, underscoring its potential for intelligent data center security management (see Table III).

TABLE III. YOLO MODEL COMPARISON (V5 VS V8 VS V9)

Model	Backbone	mAP (%)	FPS	Model Size (MB)	Notes
YOLOv5	CSPDarknet53	92.3	80	145	Good trade-off accuracy/speed
YOLOv8	CSPDarknet53-P	93.5	75	160	Improved feature extraction
YOLOv9	Transformer-based	95.2	78	155	High accuracy and real-time IoT performance

B. Ablation Study Analysis

The YOLOv9 model achieves exceptional performance through an enhanced feature extraction backbone that significantly advances earlier architectures, such as the Darknet-53 backbone used in YOLOv5. While Darknet-53 utilized residual connections to maintain gradient flow and support multi-scale feature representation, YOLOv9 introduces two key innovations: the Programmable Gradient Information (PGI) module and the Generalized Efficient Layer Aggregation Network (GELAN) backbone. GELAN combines the strengths of CSPNet and ELAN designs, enabling the model to maintain gradient consistency and capture hierarchical features across multiple network depths. This allows YOLOv9 to efficiently extract both low-level spatial patterns and high-level semantic information, improving detection accuracy for objects of diverse sizes, shapes, and complexities in IoT-based environments.

YOLOv9 employs a refined multi-scale detection strategy, featuring three detection heads operating at different resolutions. Each head utilizes an adaptive anchor-free mechanism to predict bounding boxes, confidence scores, and class probabilities with high precision. This design ensures accurate detection of objects ranging from small hardware components to large infrastructure elements in data centers. Convolutional layers are organized into optimized, computation-efficient blocks that integrate batch normalization, Sigmoid Linear Unit (SiLU) activation functions, and dynamic convolutional kernels, enhancing feature reuse, spatial understanding, and training stability.

Compared to YOLOv5, YOLOv9 further improves stability and accuracy through dynamic normalization layers, adaptive padding, and programmable gradient control. The SiLU activation function replaces Leaky ReLU, providing smoother gradient flow and better representation of complex nonlinear relationships. Adaptive padding and precise feature alignment strategies preserve spatial consistency, minimizing information loss at image boundaries and improving localization of small or partially occluded objects. Feature fusion leverages Cross-Stage Partial (CSP) connections and nearest-neighbor upsampling, combining high-resolution shallow features with semantically rich low-resolution deep features to improve detection robustness across scales. The anchor-free multi-scale detection approach eliminates reliance on pre-defined anchor boxes, allowing the model to dynamically predict bounding boxes, confidence scores, and class probabilities. This reduces computational overhead while enhancing flexibility (Table IV).

TABLE IV. ABLATION STUDY RESULTS

Ablation Component	mAP (%)	FPS	Notes
Full YOLOv9	95.2	78	Complete system with all modules.
Without Thermal Module	92.7	79	Evaluates impact of thermal imaging.
Without Motion Detection	93.0	79	Evaluates contribution of motion sensors.
YOLOv9 w/ Darknet-53 Backbone	93.5	80	Tests alternative backbone performance.
Without Data Augmentation	91.8	78	Measures effect of dataset enrichment.

The GELAN backbone, in combination with the PGI module, ensures superior representation learning, gradient stability, and efficient parameter utilization across diverse IoT scenarios. Collectively, these architectural and functional enhancements make YOLOv9 a powerful and flexible solution for real-time IoT-based data center monitoring, offering improved detection accuracy, environmental adaptability, and computational efficiency (see Fig. 1).

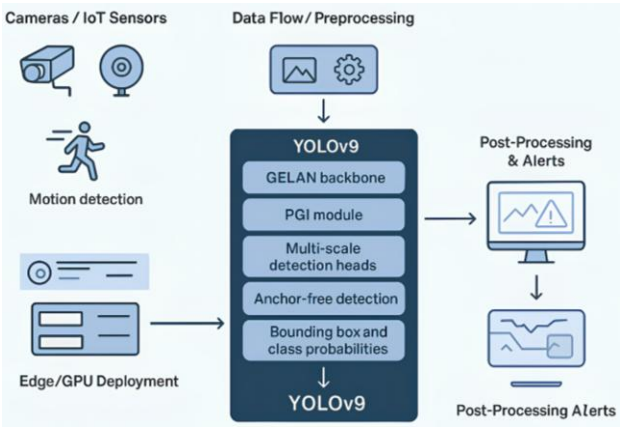

Fig. 1. YOLOv9 architecture details.

C. Environmental Robustness Analysis

Environmental robustness is a critical requirement for intelligent monitoring systems deployed in data center environments, where operational conditions are inherently dynamic. Variations in illumination, temperature, equipment density, and human activity can significantly affect visual perception and detection reliability. This section evaluates the robustness of the proposed YOLOv9-based monitoring framework under diverse environmental conditions through a series of controlled experiments.

1) Experimental design and environmental scenarios

To systematically assess robustness, a set of controlled experiments was conducted by isolating key environmental factors while keeping the model architecture and training parameters unchanged. The evaluated conditions include:

(1) Lighting variations (normal illumination, low-light conditions, and uneven artificial lighting);
(2) Thermal fluctuations (normal operating temperature versus localized overheating captured via thermal imaging);
(3) Spatial density changes (standard versus high-density equipment layouts);
(4) Human activity levels (no presence, authorized operators, and simulated unauthorized movements).

Each scenario was evaluated independently to measure the sensitivity of the model to environmental perturbations and to analyze performance degradation patterns.

2) Control experiments and baseline configuration

A baseline configuration was established using standard operating conditions: stable lighting, nominal temperature ranges, regular equipment spacing, and controlled human activity. Performance under this baseline was compared against stressed environmental conditions to quantify robustness. The same dataset splits, confidence thresholds, and inference settings were applied across all experiments to ensure fair and reproducible comparisons.

3) Evaluation metrics

Robustness was evaluated using standard object detection metrics, including precision, recall, mean Average Precision (mAP) @0.5, and inference latency. In addition, a robust score was defined to capture relative performance degradation compared to the baseline scenario. This score provides a quantitative indicator of system stability under environmental stress.

The results demonstrate that the proposed framework maintains high detection accuracy across all tested conditions, with limited performance degradation under environmental stress. The most significant impact was observed under low-light and high human activity scenarios, where visual ambiguity increases. However, the integration of multimodal inputs and data augmentation strategies mitigated performance loss, maintaining mAP above 93% in all cases.

Thermal variability had a comparatively lower impact, indicating that the fusion of RGB and thermal information enhances robustness against temperature-related disturbances. Slight increases in inference latency were observed under stressed conditions, but these remained within acceptable real-time constraints for data center monitoring.

Overall, the control experiments confirm that the proposed YOLOv9-based framework exhibits strong environmental robustness, making it suitable for deployment in real-world data center environments characterized by dynamic and unpredictable conditions (Table V).

TABLE V. ENVIRONMENTAL ROBUSTNESS CONTROL EXPERIMENTS UNDER VARYING CONDITIONS

Environmental Condition	Precision (%)	Recall (%)	mAP@0.5 (%)	Latency (ms)	Performance Drop (%)
Baseline (Control)	96.4	95.8	96.1	18.2	0.0
Low-light conditions	93.7	92.9	93.2	18.9	2.9
Uneven lighting	94.1	93.4	93.8	19.1	2.3
High thermal variance	95.2	94.6	94.9	18.6	1.2
High equipment density	94.5	93.8	94.1	19.4	2.0
High human activity	93.9	93.1	93.5	19.7	2.6

D. Case Visualization and Analysis

The YOLOv9 model introduces several architectural and functional enhancements that significantly improve upon the principles established by YOLOv5. While YOLOv5 relied on fixed padding, batch normalization, and Leaky ReLU activations for stabilization and feature learning, YOLOv9 advances these concepts with generalized normalization layers, dynamic activation functions, and programmable gradient control, enabling faster convergence and improved accuracy. To enhance training stability, YOLOv9 integrates dynamic batch normalization mechanisms that adaptively adjust normalization parameters across layers, ensuring consistent gradient propagation even in large-scale datasets. Furthermore, the SiLU activation function replaces Leaky ReLU, offering smoother gradient flow and better representation of complex nonlinear relationships. Instead of traditional fixed padding, YOLOv9 employs adaptive padding and precise feature alignment strategies that preserve spatial consistency throughout the convolutional hierarchy. This ensures minimal loss of information at image boundaries and enhances localization precision for small or partially occluded objects. For feature fusion, YOLOv9 uses efficient upsampling blocks based on nearest-neighbor interpolation and Cross-Stage Partial (CSP) connections. This mechanism allows the model to merge high-resolution features from shallow layers with semantically rich, low-resolution representations from deeper layers, improving detection robustness across varying object scales. YOLOv9 adopts an anchor-free multi-scale detection approach, replacing the fixed anchor box system of YOLOv5. This allows the model to dynamically predict bounding boxes, confidence scores, and class probabilities without relying on pre-defined anchors, thereby improving flexibility and reducing computational overhead. The detection heads operate at three distinct scales, each optimized through adaptive receptive fields to detect small, medium, and large objects with higher precision. The Generalized Efficient Layer Aggregation Network (GELAN) backbone replaces the Darknet-53 feature extractor, offering enhanced gradient flow, parameter efficiency, and deeper semantic understanding. Combined with the Programmable Gradient Information (PGI) module, YOLOv9 achieves superior representation learning and optimization stability across diverse IoT-based environments. By incorporating these innovations—dynamic normalization, adaptive padding, multi-scale fusion, and anchor-free detection—YOLOv9 delivers substantial improvements in real-time object detection performance, environmental adaptability, and computational efficiency, making it particularly well-suited for IoT-driven data center surveillance and monitoring applications.

The detection process in YOLOv9 is executed using optimized parameters and adaptive thresholding mechanisms to enhance precision and reduce false detections. Instead of relying on fixed parameters such as an IoU threshold of 0.5 and a confidence threshold of 0.5, YOLOv9 integrates dynamic threshold tuning that adjusts these values based on the scene complexity and object density. This adaptive configuration enables the model to maintain stable detection accuracy across diverse IoT-based environments. During inference, the input images are preprocessed and batched according to the selected model size and computational capacity. The batch size is dynamically managed to balance GPU utilization and latency. The class definitions are loaded from a configuration file, and the number of target classes is automatically initialized. Key hyperparameters, including the maximum output size, IoU threshold, and confidence threshold, are fine-tuned through the model's internal optimization routine. Once the model is initialized, the pre-trained YOLOv9 weights, typically stored in a file such as YOLOv9.pt, are loaded through the load_weights function. These weights encapsulate extensive feature learning obtained from large-scale datasets, ensuring robust generalization across different visual contexts. The inference engine (e.g., TensorRT or ONNX Runtime) executes the forward pass to generate bounding box coordinates, object confidence scores, and class probabilities for each detected entity. To visualize the detection results, the draw_boxes or plot_detections function is employed. This utility overlays bounding boxes and class labels on the input images, providing a clear graphical representation of the detected objects. YOLOv9 further enhances visualization with color-coded confidence mapping and multi-class annotation, enabling users to interpret results intuitively. The integration of dynamic thresholds, efficient inference optimization, and enhanced visualization tools allows YOLOv9 to deliver high-precision, real-time object detection results suitable for IoT-based monitoring systems, data center security, and autonomous decision-making applications.

In YOLOv9, the task of loading class names from a file and returning them as a usable list is performed by the load_class_names function. This function opens the specified file in read mode, retrieves its content, and separates the class labels using the split-lines method. The resulting list provides a direct mapping between detected objects and their corresponding class labels, such as "person", "server", "router", or "cooling unit". These class names are essential for interpreting YOLOv9 detection outputs. Each detected object is assigned a unique label, allowing the system to categorize and annotate objects accurately. The list of class names is subsequently used by the draw_boxes or plot_detections functions to visually annotate the detected objects on input images. Bounding boxes, class labels, and confidence scores are overlaid, providing an intuitive understanding of the detection results. This visualization is particularly useful in data centre monitoring, where precise object recognition is critical for security and operational oversight. Fig. 1 illustrates the practical application of YOLOv9 in a data centre environment. Fig. 2 shows the overall detection configuration, highlighting the integration of advanced surveillance technologies for facility monitoring and security. Fig. 3 demonstrates the detection of unauthorized personnel, providing insight into the design of secure data workspaces. Fig. 3 emphasizes perimeter monitoring and

access control, illustrating how detection of individuals near sensitive areas enhances overall safety. Fig. 4 focuses on internal monitoring and intrusion detection, showcasing the identification of personnel within the data centre and the protection of critical equipment. These examples demonstrate how YOLOv9-based object detection systems can be seamlessly integrated into data centre infrastructures to reduce security risks, enhance operational efficiency, and safeguard valuable data assets.

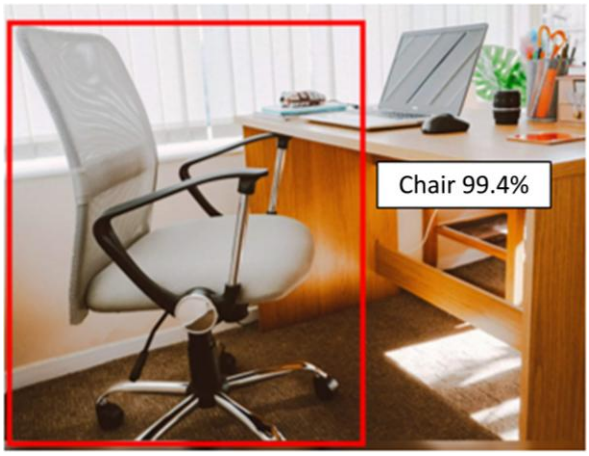

Fig. 2. Detection of the setup in a data warehouse.

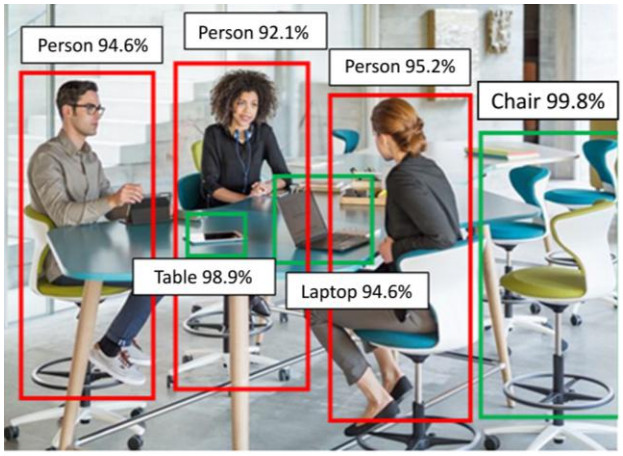

Fig. 3. Data centre security to detect intruding persons and layout of data workspace.

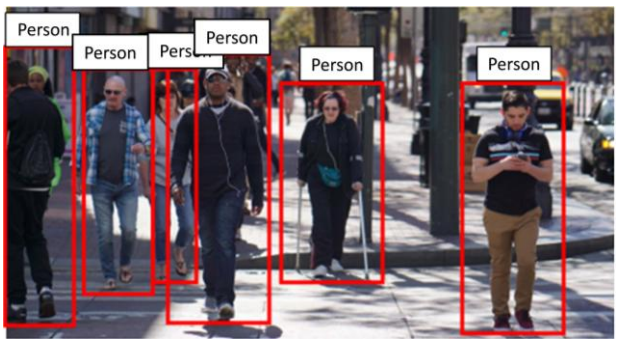

Fig. 4. Detection of people around the data centre.

Quantitative results are summarized in Table VI, showing confidence scores, and Table VII, presenting the pruning ratios after dataset partitioning, which collectively highlight the system's performance and reliability.

The function pulls the matching boxes and class labels from the boxes_dicts dictionary for each image in the input list. Drawing bounding boxes and labels for the discovered items uses the class names, confidence scores, and model size information that are provided.

TABLE VI. CONFIDENCE SCORE BASED ON RESOLUTION AFTER SET DIVISION

Images after Set Division	Resolution	Confidence Score (%)
Set 1	125×125	68.45
Set 2	250×250	73.83
Set 3	500×500	84.92

TABLE VII. PRUNING RATIO BASED ON RESOLUTION AFTER SET DIVISION

Image Set	Pruning Ratio
Set 1	0.929
Set 2	0.942
Set 3	0.998

Real-time violence detection in government surveillance systems leverages advanced computer vision and deep learning techniques to enhance public safety and operational efficiency. These systems enable rapid identification of violent incidents, allowing authorities to respond promptly and deploy resources effectively, thus minimizing harm and preventing escalation. Beyond incident response, real-time detection functions as a crime deterrent by signaling that violent actions are likely to be detected immediately. By analyzing behavioral patterns, such systems can also identify precursors to violence, providing early warnings and facilitating proactive intervention. In addition to real-time alerts, these technologies support post-incident investigation by generating reliable visual evidence for reconstructing events, identifying perpetrators, and assisting legal proceedings. Integration with emergency response networks ensures that alerts trigger the timely dispatch of personnel and resources, optimizing response times. Automation of detection processes allows for efficient resource utilization, enabling personnel to focus on critical tasks rather than manual video review. Systems can trigger proactive security measures, such as alarms, notifications, or dynamic camera adjustments, enhancing situational control and reducing potential damage. While delivering significant public safety benefits, the deployment of these systems must address privacy and data protection. Techniques such as anonymization and encryption are essential to comply with regulations while maintaining detection effectiveness.

In summary, real-time violence detection systems provide a robust framework for intelligent surveillance, combining rapid detection, early warning, proactive security actions, and post-event analysis. Their integration into government surveillance networks contributes to safer public spaces and optimized operational efficiency while maintaining compliance with privacy standards (Fig. 5).

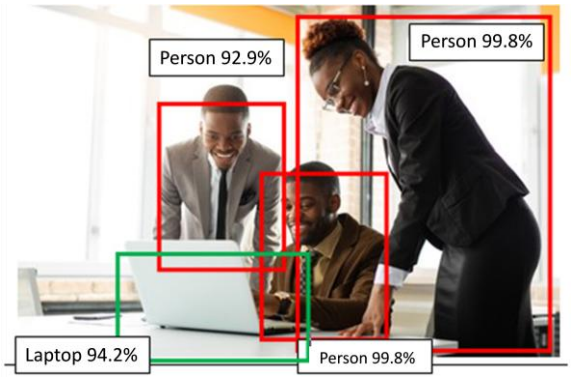

Fig. 5. Detection of people working inside the data center.

E. Error Analysis and Discussion of Limitations

A comprehensive error analysis was conducted to better understand the limitations of the proposed YOLOv9-based monitoring framework and to identify scenarios where detection performance may degrade. Analyzing misdetections and failure cases is essential for interpreting experimental results, improving system reliability, and guiding future enhancements. The most common errors observed during evaluation fall into three main categories: false positives, false negatives, and localization inaccuracies.

False positives primarily occurred in visually cluttered environments where cables, shadows, or reflective surfaces were mistakenly detected as equipment components or human presence. These errors were more frequent in zones with dense infrastructure layouts and complex lighting patterns. Although confidence threshold tuning reduced their occurrence, some ambiguity remains in highly congested visual scenes.

False negatives were mainly associated with partial occlusions or extreme viewing angles, where critical objects such as servers or cooling units were only partially visible. In addition, rapid human movement or short-duration appearances occasionally led to missed detections, particularly under low-light conditions. These cases highlight the inherent challenges of real-time monitoring in dynamic operational environments.

Localization errors were observed when bounding boxes did not accurately align with object boundaries, especially for elongated objects such as cable bundles or rack-mounted components. While these errors had a limited impact on overall detection metrics, they may affect downstream tasks such as precise equipment identification or spatial reasoning.

Despite its strong overall performance, the proposed framework presents several limitations. First, reliance on vision-based sensing makes the system sensitive to extreme environmental conditions such as severe occlusion, abrupt illumination changes, or camera misalignment. Although multimodal integration mitigates some of these effects, complete robustness under all conditions cannot be guaranteed. Second, the dataset, while diverse, may not fully capture all possible configurations and layouts encountered across different data center architectures. Variations in rack design, equipment branding, or facility-specific layouts could impact generalization performance when deploying the system in unseen environments.

Third, real-time performance is influenced by hardware constraints at the edge. While YOLOv9 achieves low inference latency, resource-limited edge devices may experience performance degradation under high data throughput or simultaneous multimodal processing. This limitation necessitates careful hardware selection and workload balancing in large-scale deployments. Finally, privacy-preserving mechanisms such as on-device inference and anonymization introduce additional computational overhead. Although this overhead remains acceptable within the evaluated setup, it may affect scalability in scenarios with many cameras or high-resolution input streams. To address these limitations, future work will focus on expanding the dataset to include more diverse data center environments and rare operational scenarios. Adaptive confidence thresholds and context-aware post-processing strategies will be explored to further reduce false detections. Additionally, lightweight model variants and dynamic resource allocation strategies will be investigated to improve scalability on constrained edge hardware. Incorporating self-supervised learning and continual learning mechanisms could further enhance long-term adaptability and robustness.

V. Conclusion

The YOLOv9 model provides a highly efficient and robust framework for IoT-based object detection in smart city environments. By combining advanced deep neural networks with state-of-the-art computer vision techniques, YOLOv9 can accurately detect and localize multiple objects within an image. The model's backbone, Darknet-53, enables effective feature extraction, while subsequent detection layers facilitate precise object recognition. To enhance detection accuracy and speed, YOLOv9 leverages batch normalization, fixed padding, and multi-scale feature maps, allowing it to handle objects of varying sizes and aspect ratios. The detection workflow begins with loading class names, which assign meaningful labels to identified objects, enabling interpretable results. These labels allow for the recognition of specific categories, such as vehicles, pedestrians, animals, and other urban entities. During inference, preprocessed images are fed into the model. YOLOv9 generates bounding box predictions, class probabilities, and confidence scores for each detected object. Low-confidence detections are filtered out, and Non-Maximum Suppression (NMS) is applied to eliminate overlapping bounding boxes by adjusting IoU and confidence thresholds, ensuring clean and accurate outputs. The visualization of detection results relies on functions such as draw_boxes, which annotate images with bounding boxes, class labels, and confidence scores. The resulting annotated images provide clear, interpretable visualizations of detected objects, supporting real-time monitoring, situational awareness, and decision-making in smart city IoT applications.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Yousef Farhaoui (Y.F.) conceived and designed the proposed IoT-enabled intelligent data center monitoring framework, supervised the overall research process, coordinated the integration of YOLO-based image analysis with IoT technologies, and contributed to the system architecture design, experimental validation, and manuscript preparation. Ahmad El Allaoui (A.E.A.) was responsible for data collection, dataset annotation, implementation of the YOLO-based object detection model, integration of motion detection and thermal imaging modules, and contributed to conducting experiments and analyzing performance results. Hamed Taherdoost (H.T.) contributed to the methodological design, performance evaluation strategy, analysis of experimental results, and provided critical insights into system scalability, optimization, and validation, and assisted in refining the research methodology and discussion. Bharat Bhushan (B.B.) contributed to the literature review, comparative analysis with related works, interpretation of results in the context of smart cities and green data center infrastructures, and assisted in improving the clarity, structure, and technical quality of the manuscript. All authors reviewed, edited, and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

REFERENCES

[1] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732–794, 2016.

[2] C. Y. Wang, I. H. Yeh, and H. Y. M. Liao, "YOLOv9: Learning what you want to learn using programmable gradient information," in *Proc. European Conference on Computer Vision*, 2024, pp. 1–21.

[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi *et al*., "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[4] J. Redmon, S. Divvala, R. Girshick *et al*., "You only look once: Unified, real-time object detection," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[5] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint, arXiv:2004.10934, 2020. https://doi.org/10.48550/arXiv.2004.10934

[6] G. Jocher *et al*. (2023). YOLOv8: Ultralytics YOLO. [Online]. Available: https://github.com/ultralytics/ultralytics

[7] Y. Farhaoui, "Artificial intelligence and IoT-enabled power electronics for renewable energy and smart grids," *IEEE Power Electronics Magazine*, vol. 12, pp. 115–118, 2026. doi: 10.1109/MPEL.2025.3624921

[8] D. Suresh, Y. Farhaoui, B. S. K. Durai, and K. Priyadarsini, "A smart Intersected Visualized Anisotropic Optimized (IntVAO) tracker for monitoring crowds and spotting abnormalities in video surveillance systems," *Journal of Visual Communication and Image*, vol. 111, 104552, 2025. doi: 10.1016/j.jvcir.2025.104552

[9] S. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6479–6488.

[10] N. M. Kumar, Y. Farhaoui, R. Vimala *et al*., "Innovative control techniques for enhancing signal quality in power applications: Mitigating electromagnetic interference," *Algorithms*, vol. 18, no. 5, 288, 2025. doi: 10.3390/a18050288

[11] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[12] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–449, 2002.

[13] T. Hassner, Y. Itcher, and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior," in *Proc. CVPR Workshops*, 2012, pp. 1–6.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735

[15] E. Breck, S. Cai, E. Nielsen *et al*., "The ML test score: A rubric for ML production readiness," in *Proc. 2017 IEEE International Conference on Big Data*, 2017, pp. 1123–1132.

[16] J. Dean and L. A. Barroso, "The tail at scale," *Communications of the ACM*, vol. 56, no. 2, pp. 74–80, 2013.

[17] N. P. Jouppi, C. Young, N. Patil *et al*., "In-datacenter performance analysis of a tensor processing unit," in *Proc. 44th Annual International Symposium on Computer Architecture*, 2017, pp. 1–12.

[18] Y. Farhaoui, "Design and implementation of an intrusion prevention system," *International Journal of Network Security*, vol. 19, no. 5, pp. 675–683, 2017. doi: 10.6633/IJNS.201709.19(5).04

[19] Y. Farhaoui, S. Ojo, L. A. Akinyemi *et al.*, "Editorial," *Big Data Mining and Analytics*, vol. 6, no. 3, pp. I–II, 2023. doi: 10.26599/BDMA.2022.9020045

[20] Y. Farhaoui, "Intrusion prevention system inspired immune systems," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 2, no. 1, pp. 168–179, 2016.

[21] Y. Farhaoui, "Big data analytics applied for control systems," in *Proc. International Conference on Advanced Information Technology, Services and Systems*, 2018, pp. 408–415. https://doi.org/10.1007/978-3-319-69137-4_36

[22] Y. Farhaoui, A. E. Allaoui, J. Rasheed, and O. Osman, "Fine-tuned object detection for mask recognition using green computing in IoT systems," *Journal of Image and Graphics United Kingdom*, vol. 13, no. 5, pp. 561–569, 2025. doi: 10.18178/joig.13.5.561-569

[23] Y. Farhaoui, "Teaching computer sciences in Morocco: An overview," *IT Professional*, vol. 19, no. 4, pp. 12–15, 2017. doi: 10.1109/MITP.2017.3051325

[24] Y. Farhaoui, "Securing a local area network by IDPS open source," *Procedia Computer Science*, vol. 110, pp. 416–421, 2017. https://doi.org/10.1016/j.procs.2017.06.106

[25] S. A. Triantafyllou, T. Sapounidis, and Y. Farhaoui, "Gamification and computational thinking in education: A systematic literature review," *Salud, Ciencia y Tecnologia-Serie de Conferencias*, vol. 3, 659, 2024. doi:10.56294/sctconf2024659

[26] P. S. Saravanan, S. Ramani, V. R. Reddy, and Y. Farhaoui, "A novel approach of privacy protection of mobile users while using location-based services applications," *Ad Hoc Networks*, vol. 1491, 103253, 2023. doi: 10.1016/j.adhoc.2023.103253

[27] R. Shamim and Y. Farhaoui, "Enhancing cloud-based machine learning models with federated learning techniques," in *Proc. The International Conference on Artificial Intelligence and Smart Environment*, 2023, pp. 594–606. doi: 10.1007/978-3-031-48573-2_85

[28] S. S. Alaoui and Y. Farhaoui, "Machine learning for early fire detection in the oasis environment," in *Proc. the International Conference on Artificial Intelligence and Smart Environment*, 2023, pp. 138–143. doi: 10.1007/978-3-031-48573-2_20

[29] N. Khouibiri, Y. Farhaoui, and A. E. Allaoui, "Design and analysis of a recommendation system based on collaborative filtering techniques for big data," *Intelligent and Converged Networks*, vol. 4, no. 4, pp. 296–304, 2023. doi: 10.23919/icn.2023.0024

[30] S. Khetavath, N. C. Sendhilkumar, P. Mukunthan *et al*., "An intelligent heuristic manta-ray foraging optimization and adaptive extreme learning machine for hand gesture image recognition," *Big Data Mining and Analytics*, vol. 6, no. 3, pp. 321–335, 2023. doi: 10.26599/BDMA.2022.9020036

[31] N. Khouibiri and Y. Farhaoui, "How can cloud bi contribute to the development of the economy of SMEs? Morocco as model", in *Proc. The International Conference on Artificial Intelligence and*

Smart Environment, 2024, pp. 149–159. doi: 10.1007/978-3-031-48465-0_20
[32] V. Veeraiah, P. Gangavathi, S. Ahamad *et al*., "Enhancement of meta verse capabilities by IoT integration," in *Proc. 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022, pp. 1493–1498.
[33] N. Gupta, S. Janani, R. Dilip *et al*., "Wearable sensors for evaluation over smart home using sequential minimization optimization-based random forest," *International Journal of Communication Networks and Information Security*, vol. 14, no. 2, pp. 179–188, 2022.
[34] D. Pudasaini and A. Abhari, "Scalable object detection, tracking and pattern recognition model using edge computing," in *Proc. 2020 Spring Simulation Conference*, 2020, pp. 1–11.
[35] S. Xie, Y. Zhou, I. Zhong *et al*., "A package auto-counting model based on tailored YOLO and DeepSort techniques," in *Proc. MATEC Web of Conferences*, 2022, vol. 355, 02054.
[36] B. K. Pandey, D. Pandey, A. Gupta *et al.*, "Secret data transmission using advanced morphological component analysis and steganography," in *Proc. Role of Data-Intensive Distributed Computing Systems in Designing Data Solutions*, 2023, pp. 21–44.
[37] V. Veeraiah, K. R. Kumar, P. L. Kumari *et al*., "Application of biometric system to enhance the security in virtual world," in *Proc. 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2023, pp. 719–723.
[38] R. Bansal, A. Gupta, R. Singh, and V. K. Nassa, "Role and impact of digital technologies in e-learning amidst covid-19 pandemic," in *Proc. 2021 Fourth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, 2021, pp. 194–202.
[39] V. Jain, S. M. Beram, V. Talukdar *et al*., "Accuracy enhancement in machine learning during blockchain based transaction classification," in *Proc. 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2022, pp. 536–540.
[40] Q. Wu and X. Nie, "Improved YOLOv10: A real-time object detection approach in complex environments," *Sensors*, vol. 25, no. 22, 6893, 2025. https://doi.org/10.3390/s25226893
[41] M. Gupta, S. Ghatak, A. Gupta, and A. L. Mukherjee, *Artificial Intelligence on Medical Data: Proceedings of International Symposium*, ISCMM 2021, Springer Nature, 2022, vol. 37.
[42] R. V. Sevilla, A. S. Alon, M. P. Melegrito *et al*., "Mask-vision: A machine vision-based inference system of face mask detection for monitoring health protocol safety," in *Proc. 2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, 2021, pp. 1–5.
[43] P. Pandiyan, R. Thangaraj, M. Subramanian *et al.*, "Real-time monitoring of social distancing with person marking and tracking system using YOLOv3 model," *International Journal of Sensor Networks*, vol, 38, no. 3, pp. 154–165, 2022.
[44] X. Zhang, T. Zhang, Y. Yang, Z. Wang, and G. Wang, "Real-time golf ball detection and tracking based on convolutional neural networks," in *Proc. 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, pp. 2808–2813.
[45] R. Dey, D. Bhattacharjee, and M. Nasipuri, "Object detection in rainy condition from video using YOLO based deep learning model," in *Proc. Advanced Computing and Systems for Security: Volume Twelve*, 2020, pp. 121–131.
[46] Q. Wang, Z. Wang, B. Li, and D. Wei, "An improved YOLOv3 object detection network for mobile augmented reality," in *Proc. 2021 IEEE 7th International Conference on Virtual Reality (ICVR)*, 2021, pp. 332–339.
[47] D. Ma, H. Fang, N. Wang, C. Zhang, J. Dong, and H. Hu, "Automatic detection and counting system for pavement cracks based on PCGAN and YOLO-MF," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22166–22178, 2022.
[48] H. Li, K. Yin, X. Ji *et al.*, "Improved YOLOV3 surveillance device object detection method based on federated learning," in *Proc. 2022 4th International Conference on Data-Driven Optimization of Complex Systems (DOCS)*, 2022, pp. 1–6.
[49] G. Kaur and S. Singh, "Violence detection in videos using deep learning: A survey," in *Proc. Advances in Information Communication Technology and Computing*, 2022, pp. 165–173.
[50] J. D. Wu, B. Y. Chen, W. J. Shyr, and F. Y. Shih, "Vehicle classification and counting system using YOLO object detection technology," *Traitement du Signal*, vol. 38, no. 4, 2021.
[51] A. Bondalapati, S. N. Bhavanam, and E. S. Reddy, "Moving object detection based on unified model," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 6, pp. 6057–6072, 2021.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).